2019

Improving Undergraduate Statistics Education: Educational Lessons From Two Pedagogical Experiments

Anushka Karkelanova
University of Kentucky, aakark2@g.uky.edu
Digital Object Identifier: https://doi.org/10.13023/etd.2020.044

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Anushka Karkelanova, Student

Dr. Xin Ma, Major Professor

Dr. Kristen Perry, Director of Graduate Studies

IMPROVING UNDERGRADUATE STATISTICS EDUCATION:
EDUCATIONAL LESSONS FROM TWO PEDAGOGICAL EXPERIMENTS

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements of the requirements for the degree of Doctor of Education in the College of Education at the University of Kentucky

By

Anushka Karkelanova

Lexington, Kentucky

Director: Dr. Xin Ma, Professor of Educational, School, and Counseling Psychology

Lexington, Kentucky

2019

ABSTRACT OF DISSERTATION

IMPROVING UNDERGRADUATE STATISTICS EDUCATION:

EDUCATIONAL LESSONS FROM TWO PEDAGOGICAL EXPERIMENTS

The ultimate goal of statistics education is to create a statistically literate society in which people can appropriately use statistical thinking. Although the need to improve the teaching of introductory statistics courses is not a new one, with increased demand on these courses, there has been constant effort to seek out better ways of teaching these courses. The University of Kentucky (UK) began a reform of its general education program in November 2005. Thinking and reasoning are the central themes of this well-designed general education curriculum.

The main goal of this dissertation is to fill in some gaps in the research literature on the teaching and learning of statistics. This dissertation includes two independent studies (experiments). The first study will examine the instructional effects of physical versus virtual manipulatives (see definitions later) on learning outcomes in introductory statistics, whereas the second study will investigate the impact of different styles in teaching statistics (inverted classroom versus traditional classroom) on learning outcomes in introductory statistics. In general, this dissertation strives to join many other reform efforts to explore instructional ways that engage students in reasoning and thinking statistically. To combat the abstract nature of probability and statistics, the use of manipulatives may represent one of the most effective strategies in the statistics classroom. There are fundamental reasons to inherently value the inverted classroom’s emphasis on activity-based learning and increased responsibility of the students to become active participants in their own learning.

The results of the first study revealed that there were no significant differences between the business as usual group who received traditional concrete manipulatives and the experimental group who received online virtual manipulatives. There -were no statistically significant interaction effects between types of manipulatives and high school ACT mathematics scores, informing the literature that ability levels neither intensify nor weaken the effects of types of manipulatives. The results of the study did not show a significant difference in GPA one year later between the experimental group and the business as usual group.

The results of the second study revealed that there were some significant differences between the business as usual group who received traditional lecture type classroom and the experimental group who received inverted. We compared all seven outcomes for the two groups: projects average, tests average, classwork, midterm attendance average, class final attendance average, midterm grade and class final grade. Students in the traditional classroom did better than students in the inverted classroom in projects average, classwork, midterm attendance average, midterm grade and class final grade. We used tree different blocks with student background variables as predictors. The first one, individual student background, is explained by age, gender and ethnicity. High school background variables is explained by high school GPA and ACT mathematics scores. The third one, university program background, is explained by university cumulative GPA and student major.

After controlling for student background variables, students in the traditional classroom did better than students in the inverted classroom in projects average, overall classwork and midterm grade. The model when controlling for student high school background variables showed that students in the traditional classroom did better than students in the inverted classroom in projects average, overall classwork and midterm grade. Finally, after controlling for student university background variables, students in the traditional classroom performed similarly to students in the inverted classroom in projects average, test average, overall classwork, midterm attendance average, class final attendance, midterm grade and class final grade. When controlling for all (i.e., student background variables, student high school background variables, and university program variables), students in the traditional classroom did better than students in the inverted classroom in midterm grade only.

The results of the study may not be generalized to the population of all undergraduate students. It also gives no indication of how the results would generalize to other content domains. Further studies may explore along these lines of inquiry regarding the effects of virtual manipulatives in comparison with concrete manipulatives and the effects of the traditional classroom in comparison with inverted. Further studies may seek some longer period of using and comparing the two teaching methods.

KEYWORDS: Statistics Education, Statistics Achievement, Student Background, Inverted classroom, Virtual Manipulatives, Manipulatives

Anushka Karkelanova

December 5, 2019

Date

ACKNOWLEDGEMENTS

I wish to express my deep gratitude to my Dissertation Chair, Dr. Xin Ma, for his kindness and hours of dedicated service on my behalf. May he be richly blessed for his selfless work and encouragement throughout the research and writing process. His and Dr. William Rayens' ongoing patient assistance was even more critical in developing the project and preparing the dissertation itself. Dr. Rayens, I thank you for your guidance, support, and mentorship for the last six years. I appreciate how you have always been an advocate of my work, and for the ways you always challenge my thinking about statistics education. I could not have finished this dissertation without your support and flexibility along the way.

In addition, I also wish to thank Dr. Kenneth Tyler and Dr. Garry Swan as members of my committee for their insights on sharpening and understanding the aspects of statistics education research. Thanks, as well, to Dr. Jungmin Lee for her willingness to engage in this process as an outside examiner.

In addition to the essential support provided by the suite of five academic experts above, I would not have finished this work without the remarkable patience of my husband, Entcho, who stood with me through this extended process. Entcho, thank you for your endless love and making this endeavor a possibility for me. Without your constant support and encouragement, I would have never made it to this point. Thank you for walking alongside me every step of this journey and for all of the sacrifices you endured along the way. I could not have completed this dissertation without your unceasing support and love. All my family, husband and children Eric and Petya have been remarkably supportive.

Further, I am very grateful to fellow faculty and staff at the University of Kentucky (especially the Department of Statistics) and beyond for the sharpening you have provided. It has truly been a privilege to work with each of you. I have learned a great deal from you; I am grateful for your feedback that has challenged my thinking and improved my work. I am particularly grateful to Melissa Pittard, Amanda Ellis, Arnold Stromberg, Katherine Thompson and many others who have challenged my thinking and helped me to become a better educator and scholar. I could not have survived this dissertation journey without the support of my friends and colleagues. You have enriched my life, our times together walking, drinking coffee, and discussing teaching were always a source of inspiration and encouragement to me. I thank you for all of the ways that you have supported me along this journey.

Also, I must acknowledge with profound gratitude the thousands of students with whom I have been privileged to learn over the past decades. I believe that I have learned more from trying to help you learn than I have ever learned on my own.

And last, but certainly not least, I thank God for bringing amazing people to my life, and for giving me the strength and perspective to persevere and to enjoy this process.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1: STATEMENT OF THE PROBLEM

General Background

The ultimate goal of statistics education is to create a statistically literate society in which people can appropriately use statistical thinking (Kettenring, Lindsay, & Siegmund, 2004; Schau, 2003). Suggesting that statistics involves distinctive and powerful ways of thinking, Moore (1998) stated that "Statistics is a general intellectual method that applies wherever data, variation, and chance appear" (p. 1254). Because the study of statistics provides students with tools and ideas that allow them to react intelligently to quantitative information in the world around them, every high school graduate should be able to use sound statistical reasoning to intelligently cope with the requirements of citizenship, employment, and family and to be prepared for a healthy, happy, and productive life (American Statistical Association, 2005). One of the goals of statistics education at all levels is to develop statistical literacy and statistical skills in problem solving, data analysis, and data communication, as opposed to merely imparting computational procedures (Gal, 2005; Moore, 1990; National Council of Teachers of Mathematics [NCTM], 2000).

The National Scene

This dissertation research comes at a historical time when there is a strong emphasis on the need to improve students' ability to think statistically at all educational levels. Statistical reasoning, considered a powerful and important foundation for future understanding of probability and statistics, has become a key part of the mainstream school mathematics curriculum in the United States, often referred to as the Common Core State Standards for Mathematics (CCSSM) (Cohen, 2012).

CCSSM came out of serious disagreement among parents, mathematicians, and mathematics educators on policies and practices in mathematics education (Klein, 2003), a fervent emphasis of NCTM (2006, 2009) as its key vision for the critical roles of reasoning, communication, connections, and problem solving in mathematics education, and a strong belief that a set of common rigorous standards has the best chance of addressing a major deficit of public education—namely, that students are not provided with the knowledge and skills that they need to succeed (American Diploma Project, 2004). In particular, NCTM believes that organizing mathematics curriculum around its focal points can provide students with a connected, coherent, ever-expanding body of mathematical knowledge and an awareness of the unique ways of mathematical thinking. This vision of NCTM contributed in a major way to the creation of the 2010 CCSSM, which aims to provide more clarity on what students are expected to learn so as to make mathematics education more consistent across states and to guide teachers and parents in preparing students for the challenges of the workplace or postsecondary education.

The mission statement claims that the standards are designed to be robust and relevant to the real world, reflecting knowledge and skills that young people need for success in college and career, with the ultimate goal of preparing students to compete successfully in a global economy. The standards are made clear, understandable, and consistent: Include rigorous content and application of knowledge through high-order skills, be informed by educational policies and practices in top-performing countries in international comparative studies, and align with college and workplace expectations. Students are expected to succeed in entry-level, credit-bearing academic college courses and in workforce training programs typical of the global economy.

CCSSM includes content standards that strongly emphasize statistical reasoning, called *Measurement and Data* at the elementary school level and *Statistics and Probability*" at the secondary school level. A considerable amount of new curricular and instructional materials for statistics has been produced as a part of elementary and secondary mathematics with increased emphases on such activities as locating and processing quantitative information, collecting data, interpreting data and drawing inferences, and making predictions from data (CCSSM, 2010). There is a growing movement to introduce concepts of statistics and probability into the elementary and secondary mathematics curriculum, and there are calls for teaching statistics and probability in a deeper and different way than has been done (NCTM, 2000; CCSSM, 2010).

Educational reforms in K-12 mathematics education are creating considerable impact on undergraduate statistics education at the college level. Competence in statistical concepts is now valued as much as technical skills for all students (Rumsey, 2002). This broadening of what students *really need* from statistics has led to fundamental reforms in curriculum and instruction, not only in K-12 classrooms but also in college classrooms. In fact, changes in content and pedagogy (particularly concerning introductory statistics courses) have been constantly made as part of a reform effort dating back to the early 1990s (e.g., Cobb, 1992; Moore, 1997). Although the need to improve the teaching of introductory statistics courses is not a new one, with increased demand on these courses, there has been constant effort to seek out better ways of teaching these courses (e.g., Garfield, Hogg, Schau, & Whittinghill, 2002; Lindsay et al., 2004). The current emphasis is more on "awareness of data in everyday life" that prepares students "for a career in today's 'age of information'" (Rumsey, 2002, p. 2). A careful balance of content, pedagogy, and technology (Moore, 1997; Shaughnessy, 2007) helps introductory statistics courses "move beyond the 'what' of statistics to the 'how' and 'why' of statistics" (Rumsey, 2002, p. 7).

The UK (University of Kentucky) Scene

In recent years, many statisticians have become involved in the ongoing reform of the teaching of introductory statistics (e.g., Garfield & Ben-Zvi, 2008), and the National Science Foundation has funded numerous projects in promotion of this reform (e.g., Garfield et al., 2002). Moore (1997) emphasized that introductory statistics education should take place in a new social context because the changing nature of statistics as a discipline demands strong synergies among content, pedagogy, and technology (see also Garfield, 2003). The changes in content are characterized by more data analysis and less theoretical probability, the changes in pedagogy are of fewer lectures and more active learning, and the change in technology emphasizes the use of modern computing technology for data analysis and simulation.

The University of Kentucky (UK) began a reform of its general education program in November 2005 and formally implemented the new General Education Program in May 2009 (often referred to as UKCore) (see http://www.uky.edu/ukcore/). UKCore strongly emphasizes skills such as critical thinking, reasoning, writing, ethics, and global understanding so as to prepare students to compete in a global marketplace, to participate in democratic self-governance, and to live a well-intentioned and meaningful life. Thinking and reasoning are the central themes of this well-designed general education curriculum that allows students to recognize the value of critical thinking, gives them the necessary skills to reason (i.e., analyze) information critically, and offers them opportunities to witness firsthand how scholars and experts struggle to make sense out of complex problems. Overall, UKCore strives to shift graduates from a culture of actions based on opinions to a culture of actions based on evidence-based reasoning.

Specifically, UKCore is anchored by a set of four primary learning outcomes, one of which is *Quantitative Reasoning*. The expectations, according to UKCore, are:

> Students will demonstrate an understanding of and ability to employ methods of quantitative reasoning. Students will (a) demonstrate how fundamental elements of mathematical, logical and statistical knowledge are applied to solve real-world problems; and (b) explain the sense in which an important source of uncertainty in many everyday decisions is addressed by statistical science, and appraise the efficacy of statistical arguments that are reported for general consumption.

In sum, the stakes for reform in mathematics and statistics education are high not only for public schools but also for colleges and universities (Klein, 2003).

Importance of Statistics education

Many research studies over the past several decades indicate that most students and adults cannot think statistically about important issues that affect their lives (Garfield & Ben-Zvi, in press), even though their lives are increasingly governed by numbers (Moore, 1997). Konold and Higgins (2003) asserted that without sufficient statistical knowledge, it is difficult for today's citizens to have an informed opinion and participate in social and political debates concerning environment, health, education, and so on. Tishkovskaya and Lancaster (2012) argued that our society has entered into an age of information where the "information explosion" is creating a critical need for statistically educated citizens— people who need to be statistically literate not only in their workplace but also in their everyday life.

As more and more academic departments realize the importance of statistical thinking in their own disciplines, enrollments in statistics courses at the college level continue to grow (Scheaffer & Stasney, 2004). Indeed, it is widely recognized that statistics is one of the most important quantitative subjects in learning any university curriculum (Watson, 2006). As the value of statistical thinking and statistical reasoning has become more widely recognized, enrollments in statistics courses at the college level have begun to grow (Scheaffer & Stasney, 2004), and an ever-increasing number of students are taking courses in statistics to satisfy the common quantitative literacy requirement for graduation at their respective undergraduate institutions. For example, an estimated 260,000 undergraduate students in the United States enrolled in a statistics course in 2005, an increase of more than 40,000 students from 1995 (Lutzer, Rodi, Kirkman, & Maxwell, 2007). This number is likely an underestimate as it is based on enrollment in courses offered by mathematics and statistics departments and does not count students who take statistics courses in other departments (Dupuis et al., 2012). Based on website information from the National Center for Educational Statistics (NCES), enrollment in statistics courses in degree-granting institutions increased 37 percent between 2000 and 2010, from 15.3 million students to 21.0 million students.

Addressing the need to improve students' ability to think statistically, schools are making statistical reasoning a critical part of the mainstream mathematics curriculum around the world (e.g., Australian Education Council, 1994; Batanero, Burill, & Reading, 2011; Curriculum Corporation, 2006; Dani & Joan, 2004; Department for Education and Employment, 1999; Gal, 2002; Ministry of Education, 1992; National Council of Teachers for Mathematics, 2000). According to the 2007 Guidelines for Assessment and Instruction in Statistics Education, statistics has become a key component of the school mathematics curriculum in less than a quarter of a century, responding to the data richness of the society in the information age and taking advantage of the advancement in technology and modern methods of data analysis. As a result, statistical concepts are being introduced as early as elementary school. NCTM (2000) is among the most vocal for the idea that improved statistics education must begin as early as possible at the school level. According to the 2006 College Board Standards for College Success, topic areas of *Data* *and Variation* and *Chance, Fairness, and Risk* are "central to the knowledge and skills developed in the middle-school and high-school years" (p. 4). CCSSM (2010) prescribes topics of probability and statistics at each grade level. The emerging quantitative literacy movement (many ideas are statistical in nature) calls for greater emphasis on practical quantitative skills that assure success for high-school graduates in life and work (e.g., Steen, 2001). Overall, statistics education is critical in today's data-rich economy because it can promote the "must-have" competencies essential to "thrive in the modern world" (Franklin et al., 2007, p. 4).

Challenging Issues in Statistics Education

Research literature is full of students' inabilities to understand statistical concepts and procedures, a strong indication of the need for reform in statistics education. Research has identified misconceptions regarding correlation and causality (e.g., Kahneman & Tversky, 2000), conditional probability (e.g., Falk & Greenbaum, 1995; Garfield, 2003; Tarr & Lannin, 2005), independence (e.g., Tarr & Lannin,2005), randomness (e.g., Falk & Greenbaum, 1995; Fichbein & Schnarch, 1997; Konold, 1991), the Law of Large Numbers (e.g., Garfield & Ben-Zvi, 2008), and weighted averages (e.g., Reed & Jazo, 2002; Shaughnessy, 2007). In fact, inappropriate reasoning about statistical ideas is widespread and persistent at all age levels (even among some experienced researchers) (Garfield, 2002; Watson, 2013).

Cognitive Challenges

Students often consider statistics as the worst course they take while in college (Hogg, 1991). They found that the concepts of probability and statistics are very difficult to learn and often conflict with many of their own beliefs and intuitions about data and chances (Garfield & Ahlgren, 1988; Shaughnessy, 1992). According to Perney and Ravid (1991), statistics courses are viewed by most college students as a roadblock to obtaining their degrees, and students often delay taking their statistics courses until the end of their programs. Statistics is difficult not only for undergraduate students but also for graduate students in many applied fields (e.g., social sciences) (Berk & Nanda, 1998; Davis, 2003; Perepiczka, Chandler, & Becerra, 2011; Schau, Stevens, Dauphinee, & Del Vecchio, 1995). Indeed, the methods of statistics have historically been viewed by many students as difficult to understand and unpleasant to learn (Garfield & Ben-Zvi, 2007).

Ben-Zvi and Garfield (2004) discussed some of the reasons that explain why statistics is a challenging discipline to learn and to teach. First, many statistical ideas and rules are complex, difficult, and even counterintuitive so as to discourage students to engage in the learning of statistics. Second, many students have difficulty with the underlying mathematics (e.g., fractions, decimals, proportional relationship, algebraic manipulation), which interferes with the learning of statistical concepts and procedures. Third, the context in many statistical problems tends to mislead students to rely on experiences and often faulty intuitions to produce a solution rather than select an appropriate statistical procedure and rely on data-generated evidence. A fourth reason is that students equate statistics with mathematics and expect the focus to be on numbers, computations, and formulas, all leading to just one correct answer. Finally, inadequate experiences fail to prepare students for the massiveness of data, the different possible interpretations based on different assumptions, and the extensive reliance on communication skills.

Ramsey (1999) emphasized that statistics educators must understand the unique nature of the discipline and be willing to recognize the implications of that uniqueness for the teaching of statistics. According to Ramsey, the first source of difficulty comes from the fact that probability and statistics are essentially acausal. The shift from disciplines with pervasive causal interpretation to one that is inherently acausal represents a major fundamental paradigm shift in viewpoint that cannot be merely dismissed as an alternative explanation. The second source of difficulty is due to the fact that statistical reasoning is very abstract and quite foreign to the average student, even with the attempt to relate probability and statistics to observable events (i.e., the connection between theory and observation is not easily established). The final source of difficulty comes from students' attempts to relate statistical reasoning to physical cognate disciplines such as physics, chemistry, biology, and economics. Because statistics is acausal and the cognate disciplines are inherently causal, the link between the two is difficult for the student to fathom.

The famous statistician John Tukey believed that statistics is more of a science than it is a branch of mathematics. He pointed out that it is sufficient for a mathematical theorem to be elegant if it is beautiful and true, but statistics is held to an additional standard imposed by science. Velleman (2008) explained this point well by saying that a statistical model for data, no matter how elegantly and correctly derived, must be discarded or revised if it does not fit the data or if it fails to fit new or better data when available. Huxley (1893) referred to this as "the great tragedy of science" (i.e., the slaying of a beautiful hypothesis by an ugly fact) (p. 244). Overall, De Veaux (2008) argued that much of the beauty of mathematics stems from its axiomatic structure and logical development and, in fact, this structure dictates the order in which any mathematical material is taught and ensures that any mathematics course is self-contained. Unfortunately, this course design principle does not work in statistics according to these authors.

Affective Challenges

Gal, Ginsburg, and Schau (1997) added negative affective dispositions as another reason for students' difficulty in learning statistics. According to these authors, students' attitudes and beliefs regarding statistics deserve special attention for three major reasons. First, students' attitudes and beliefs toward statistics influence heavily the teaching and learning process in statistics education. Second, they influence students' statistical behaviors after they leave the classroom (i.e., in the real world). Finally, they play a major role in influencing whether or not students choose to enroll in statistics courses later on. Gal et al. (1997) went on to provide a list of beliefs that deserves consideration by those involved in statistics education. Some beliefs concern the extent to which statistics is part of mathematics or requires mathematical skills (so that negative attitudes and beliefs concerning mathematics are transferred to statistics); others center on the uncertainty of what should happen or transpire in a statistics classroom (i.e., expectations as to the culture of a statistics classroom); and still others address the usefulness or value of statistics in one's future life or career and the lack of self-confidence among students of statistics. Attitudes and beliefs concerning statistics represent a summation of experiences over time in the context of learning statistics (and mathematics). Students' negative attitudes toward statistics are an influential contributor to the low performance of students in statistics courses (Araki & Schultz, 1995; Cashin, & Elmore, 2005; Harvey & Oswald, 2000; Hilton & Schau, 2004; Mills, 2004; Mvududu, 2003; Schulz & Koshino, 1998; Waters & D'Andrea, 2002).

Pedagogical Challenges

The way that statistics has been taught also contributes to the fact that students in general consider statistics difficult to learn. Moore (1992) called for a shift from the traditional view of teaching statistics as a mathematical topic to a new view that distinguishes between mathematics and statistics as separate disciplines. Specifically, Moore (1992) argued that statistics is a mathematical science but not a branch of mathematics and has clearly emerged as a discipline in its own right with characteristic modes of thinking that are fundamentally different from any mathematical theory. Statistical theories are relative and not straightforward, with arguments based not on logics-driven consequences but on data-driven inferences (Gattuso, 2006). Hughes-Hallett (2001) also made a distinction between statistical (quantitative) literacy and mathematical knowledge, arguing that mathematical knowledge asks students to rise above context, while quantitative literacy asks students to stay within context.

Moreover, the teaching emphasis is often placed on the computation of statistical information instead of the development of an "authentic data analysis point of view" (Cobb, 1999, p. 5). Velleman (2008) asserted that statistics education ignores the guidance for students to make personally responsible statistical judgments based on a good appreciation of the role of ethics in statistics. The principle guiding statistical judgments, which is the honest search for truth about the world, should have a central place in statistics courses. Introductory statistics courses fail to recognize a common-sense approach based on examples and experiences in life (De Veaux, 2008). A student in calculus is not required to comment on whether a question makes sense and assumptions are satisfied, to evaluate the consequences of an answer, or to communicate the answer to a general audience without sufficient scientific background; however, all of these are required of students in introductory statistics courses (De Veaux, 2008).

Potential Solutions

Separation of Statistics from Mathematics

Clarification of the differences between statistics and mathematics—including the role of mathematics in statistics education—is one key solution to the challenges that statistics education faces at the college level. Apart from what has been said so much earlier, according to Moore and Cobb (1997),

> Statistics is a methodological discipline. It exists not for itself, but rather to offer to other fields of study a coherent set of ideas and tools for dealing with data. The need for such a discipline arises from the omnipresence of variability. (p. 801)

A major objective of statistics education is to help students develop statistical reasoning which, in large part, must deal with the omnipresence of variability. Statistical problem solving and decision making depend on understanding, quantifying, and explaining the variability in data. It is this focus on variability in data that sets apart statistics from mathematics.

Both Cobb (1992) and Moore (1997) concluded that the difference between statistics and mathematics has profound implications for teaching. Specifically, it is not enough to help students understand the mathematical theory behind a statistical theory; statistics teachers must also provide a ready supply of real illustrations and know how to use them to involve students in the development of their critical judgment. In mathematics, where applied context is not important, improvised examples often work well; while in statistics, improvised examples do not work well because they do not provide authentic interplays between pattern and context (Cobb, 1992; Moore, 1997). In addition, the reform in considering statistics fundamentally different from mathematics may help prevent mathematics anxiety caused by negative experiences in mathematics, which then may transfer into statistics education, given that statistics anxiety is correlated with mathematics anxiety (Gal et al., 1997; Zeidner, 2011).

Shift in Content and Pedagogy

With the separation of statistics from mathematics, statistics educators are still trying to fully understand the challenges and difficulties in teaching and learning statistics as a unique discipline. Reforms in statistics education is ongoing. For example, improving instructional materials and methods, enhancing technology, and developing alternative assessment methods have been offered as ways to reform statistics education at both school and college levels (e.g., Chance, 2005; Gal & Garfield, 2007; Garfield, 2010; Lovett & Greenhouse, 2000). One predominant reform movement at all educational levels advocates the shift of focus in content and pedagogy from computation and procedures to statistical thinking and reasoning (Garfield & Gal, 1999).

Statistics educators over the last decade have called for the development of statistical literacy and interpretive skills as the universal goals of statistics education (e.g., Del Mas, 2002; Rumsey, 2002). As early as the1990s, many statisticians started to become involved with reform movements in statistics education with the support of the National Science Foundation (Cobb, 1993). Moore (1997a) described many changes in content (e.g., more data analysis, less probability theory) and pedagogy (e.g., fewer lectures, more activities) (see also Garfield, 1995; Hoaglin & Moore, 1992). Many statisticians incorporated technology into statistics courses, particularly introductory statistics courses (e.g., for data analysis and simulation) (e.g., Chance, Ben-Zvi, Garfield, & Medina, 2007; Lock, 2000; Moore, 1997a; Seymour, 2002; Velleman & Moore, 1996).

Students' fears remain an issue for many statistics educators (Zinn & Smiley, 2003). To combat this (see Baloglu, 2003), some statisticians have attempted to address these fears in their textbooks, with titles such as *Statistics Without Tears* (Rowntree, 2004), *Statistics for the Terrified* (Kranzler, 2007), and *Statistics for People Who Think They Hate Statistics* (Salkind, 2012).

Part of this reform seeks for better alignment of instruction with important learning goals and assessments (Garfield & Gal, 1999). Assessment as a way to inform statistics educators for instructional purposes and students for progress reports, either formative or summative, is of great interest to statisticians (Mills, 2002). In fact, innovative methods of assessments are abundant (even though most of them have not been tested for pedagogical merits) (Webb, 1997). For example, calling for alternative assessment methods, Schwartz (1995) argued that traditional forms of assessment are not aligned with current curricular and instructional goals, are too narrow to provide sufficient information about student learning, and are inadequate for evaluating student understanding or promoting successful learning outcomes.

Business Involvement

The business sector has also joined forces to improve the teaching and learning of statistics. The integration of computers into statistics education has led to increased accessibility for undergraduate students and an increase in the development of more user-friendly statistics packages (e.g., SAS, SPSS, and MINITAB) (Mills, 2002). Students can now actively involve themselves in data analysis as a way to obtain a deeper understanding of statistical concepts and procedures (Brakke, Wilson, & Bradley, 2007; Garfield et al., 2002; Giesbrecht, 1996; Gratz, Volpe, & Kind, 1993; Hubbard, 1992; Marasinghe, Meeker, Cook, & Shin, 1996; McBride, 1996; Mills, 2002; Mittag, 1992; Packard, Holmes, & Fortune, 1993; Hulsizer & Woolf, 2009; Sullivan, 1993; Triola, Goodman, LaBute, Law & MacKay, 2006; Velleman & Moore, 1996). With this increasing use of technology, however, research becomes necessary to understand the effect of using technology in statistics education on student learning in statistics.

Goals of This Dissertation

Despite a growing body of research related to the teaching and learning of statistics at all educational levels, few direct connections have been established between research and practice (Garfield & Zvi, 2009). Although educational research has long been interested in the assessment of statistics anxiety and attitudes toward statistics as well as some other factors (e.g., mathematical background, motivation to learn) that predict student achievement in statistics (e.g., Garfield & Ben-Zvi, 2007), only recently have researchers started to investigate the understanding and reasoning of students concerning critical statistical concepts. Researchers are particularly interested in studying how these concepts can be developed through a carefully planned sequence of learning activities and how to implement this strategy effectively in the classroom (Garfield & Zvi, 2008). Obviously, to address this issue, empirical studies, particularly experiments in real educational settings such as a university classroom, are needed.

The main goal of this dissertation is to fill in some gaps in the research literature on the teaching and learning of statistics. By nature, this dissertation joins the reform effort of shift in content and pedagogy as discussed earlier. To promote the link between research and practice, educational experiments are used to examine the effects on learning outcomes of different instructional practices in statistics education, in particular the use of different types of manipulatives and different styles of instruction. Specifically, this dissertation includes two independent studies (experiments). The first study will examine the instructional effects of physical versus virtual manipulatives (see definitions later) on learning outcomes in introductory statistics, whereas the second study will investigate the impact of different styles in teaching statistics (inverted classroom versus traditional classroom) (see definitions later) on learning outcomes in introductory statistics. Some important student factors (e.g., prior ability) and course structure factors (e.g., availability of extra credits) also will be brought into the equation to examine whether they are capable of enhancing these treatment or intervention effects. The results of these studies will improve undergraduate statistical education and provide meaningful links between research and practice. In general, this dissertation strives to join many other reform efforts to explore instructional ways that engage students in reasoning and thinking statistically.

Definition of Terms

The first study will examine the instructional effects of physical versus virtual manipulatives on learning outcomes in introductory statistics courses. Physical (traditional) manipulatives refer to a set of concrete materials that can be physically manipulated by hands.

Virtual manipulatives refer to a set of imagines that can be electronically manipulated on a computer screen.

The second study will investigate the impact of different styles in teaching statistics (inverted classroom versus traditional classroom) on learning outcomes in introductory statistics courses. Inverted classroom refers to the instructional practice where events that traditionally take place inside of the classroom now take place outside of the classroom and vice versa. For example, students in the inverted condition may be required to watch video lectures before coming to class. This typically happens in the classroom but now becomes their homework. When students come to class, they may complete activities that are designed to help them engage in discovery learning of the content that they have already experienced by watching the videos. This is usually what students do independently after class, but now students interact with each other and the instructor in class as they work to deepen their understanding.

Traditional classroom refers to statistics classes that are taught using the traditional teaching method. Typically, students come two times a week to a classroom and listen to a lecture on certain statistical content. Often, these traditional lectures are heavily content driven, where the instructor introduces statistical concepts and then works though examples that apply those concepts. During the lectures, students may have opportunities to ask questions and answer questions from the instructor related to the content discussed. It is possible for lectures in a traditional classroom to be presented as interactively as possible.

Both studies will employ cognitive and affective measures. Cognitive learning outcomes are defined as students' academic performance in terms of (measured through) completeness of assignments, one or two specific major projects, and tests and final exams. In general, tests and exams contain both multiple-choice questions and open-ended questions. Students are usually given study guides with answers before tests and exams.

Affective outcomes (measures) include students' attitude toward statistics (interest, utility, and motivation) and confidence (anxiety) in learning statistics. Attitude, in general, is defined as "an individual's disposition to respond favorably or unfavorably to … any … discriminable aspect of the individual's world" (Ajzen, 1989, p. 241). Students' attitude toward statistics refers to students' general impression (i.e., positive or negative feelings) toward the discipline of statistics in terms of its relevance, value, and difficulty as well as the way in which self is perceived in light of the practice (e.g., learning) of statistics (see Thurstone, 1970). Such a conception considers attitude toward statistics as a multidimensional construct of interest, utility, motivation, confidence, and anxiety in the practice of statistics (see Organization for Economic Cooperation and Development, 2010). Interest refers to the level of enjoyment in the practice of statistics (e.g., liking or disliking statistics); utility refers to the usefulness, relevance, and value of statistics in life (i.e., personal and professional); and motivation refers to the amount of effort that a student is inspired to spend on the practice of statistics (Emmioglu & Capa-Aydin, 2012; Hood, Creed, & Neumann, 2012; Petocz & Newbery, 2010; Ramirez, Schau, & Emmioglu, 2012). Confidence refers to the perception of self-competence in the handling of statistical knowledge and skills in an intellectual manner (Emmioglu & Capa-Aydin, 2012; Ramirez et al., 2012), while anxiety refers to the feelings of apprehension and fear of statistics often as a result of repeated failures in the practice of statistics (Williams, 2013).

Significance of This Dissertation

This dissertation comes at a time when there is an unprecedented interest not only in taking statistics courses, but also in the reform of statistics education. Perhaps there is no other discipline that has seen a more fluid and more dynamic instructional climate in the last fifteen years than statistical science. In spite of all the interest in statistics courses and the efforts focused on pedagogical reforms in statistics instruction, for the most part there is only anecdotal evidence regarding the effectiveness of these reforms. In other words, many reform efforts hold substantial promises but have largely gone untested.

Educator and University of California (Berkeley) Professor K. P. Cross, in her 2005 paper from the Center for Studies in Higher Education, asserted that “From the instruction that we provide, to the intellectual climate that we create, to the policy decisions that we make—all should start with the question, ‘But will it improve students’ learning?’” (p. 2). This dissertation strives to address this important and challenging question by focusing on learning outcomes of undergraduate students in introductory statistics courses in relation to educational interventions (i.e., virtual versus physical manipulatives in the first study and inverted versus traditional classroom environment in the second study).

To combat the abstract nature of probability and statistics, the use of manipulatives may represent one of the most effective strategies in the statistics classroom. Manipulatives enhance the abilities of students at all levels to statistically reason and communicate, and the valuable time spent on manipulatives can also sustain long-term effects on building students’ confidence in learning statistics and deepening their statistical understanding (Shaw, 2002).

There are fundamental reasons to inherently value the inverted classroom’s emphasis on activity-based learning and increased responsibility of the students to become active participants in their own learning. What hasn’t been adequately studied is whether and how much the inverted classroom actually has a positive effect on the cognitive and affective outcomes of students. At UK, approximately 4,000 undergraduates are taught per calendar year in classrooms employing inverted statistical reasoning. A controlled experiment will produce inferential and descriptive statistical evidence either supporting the efficacy of the inverted classroom or failing to support said efficacy.

Achieving the objectives of both studies will add substantially to the limited knowledge base regarding the effectiveness of innovative reforms in shift of content and pedagogy in undergraduate statistics education. With the rising enthusiasm for educational reform in statistics education, this dissertation will provide timely insight into the effectiveness of some educational practices in undergraduate statistics education nationwide and identify factors that facilitate or hinder this effectiveness. The intellectual merit of this dissertation is both evident and substantial.

Educational reforms are becoming popular not only in statistics education but also in education of other disciplines where a passive classroom environment is no longer satisfactory to either educators or students. This dissertation therefore has a broader intellectual impact throughout higher education, in particular undergraduate education (even pre-postsecondary education). Findings from this dissertation can meaningfully inform educators in other disciplines, assisting them in the reform of their own particular conceptualizations and implementations of innovative instructions.

Finally, this dissertation will include both cognitive and affective outcome measures. This inclusion will allow this dissertation to gain a comprehensive understanding of the effects of innovative instructions. In particular, the importance of student affect in the learning of mathematics and science has been explicitly recognized and emphasized as many professional organizations (e.g., National Council of Teachers of Mathematics) advocate strongly for the improvement of affective outcomes of student learning (e.g., attitude and confidence). This dissertation will examine comprehensively the effects of innovative instruction and factors that facilitate or hinder these effects.

Overall, this dissertation will produce seminal experimental results with the potential to inform the design and implementation of inverted instruction and usage of manipulatives in the near future in statistics education and beyond.

Structure of This Dissertation

Chapters 2 and 3 are independent, self-contained chapters that document and report two independent experimental studies, both with the goal of promoting the link between research and practice. Chapter 2 covers the first educational experiment. Specifically, it will examine the instructional effects of physical versus virtual manipulatives on cognitive and affective learning outcomes in introductory statistics courses. Meanwhile, based on available data, individual and institutional factors that promote or hinder the instructional effects also will be examined. Chapter 3 covers another educational experiment, a controlled educational experiment that will investigate the impact of different instructional styles in teaching introductory statistics (i.e., inverted classroom versus traditional classroom) on cognitive and affective outcomes in statistics of undergraduate students. Specifically, inverted classroom will be considered a treatment or intervention, and the treatment effects will be assessed in comparison to traditional classroom, which will be used as a control group. Meanwhile, this experiment will collect some information on individual and institutional factors so as to examine their mediations of the treatment effects. Chapter 4 incorporates these independent studies for summaries of major findings, revisits to the research literature, implications for educational practices, limitations of this dissertation, and suggestions for future research.

CHAPTER 2

The Effects of Virtual Manipulatives on Statistics Achievement of Undergraduate Students

The purpose of this study is to fill in some gaps in the research literature on effective teaching and learning of statistics by evaluating the effectiveness of virtual manipulatives on the learning outcomes of undergraduate students in comparison with traditional manipulatives. To combat the abstract nature of probability and statistics, the use of manipulatives is one of the most effective strategies. Virtual manipulatives are technology-based innovations in statistics education designed to provide easy access to manipulatives.

Educational experiment is used to examine the instructional effects of virtual manipulatives versus physical manipulatives on the learning outcomes in introductory statistics for undergraduate students. In this posttest-only experiment, one group of undergraduate students who were enrolled in introductory statistics used traditional concrete manipulatives for learning statistics (the business as usual or BAU group), while the other group of undergraduate students enrolled in the same course used online virtual manipulatives for learning the same content (the experimental or EXP group). After one semester, undergraduate students were compared on their course average scores. Specifically, this study attempts to address the following research questions:

1. Are there any differences between the use of virtual manipulatives and physical manipulatives in the learning outcomes of introductory statistics for undergraduate students?

2. Are there any important student background variables that enhance the effects of virtual manipulatives on the learning outcomes of introductory statistics for undergraduate students?

Review of Literature

Background

Currently, researchers and statistics educators are seeking to understand the challenges and identify effective ways to overcome the difficulties in learning and teaching statistics so that improved instructional methods and materials, enhanced technology and alternative assessment methods may be used with students learning statistics at the pre-college and college level. The question of how a student best learns statistics has been heavily considered in articles on statistics teaching (e.g., Chance, 2005; Lovett & Greenhouse, 2000), and has focused mainly on instructional content or methods. In terms of instructional content, many statisticians, including Bradstreet (1996) and Cobb (1991), are convinced that an introductory statistics course should emphasize data analysis over mathematical technique and concepts over formulas. Hogg (1991) stressed that statistics should not be presented as a mathematics course at all. Statistics should emphasize statistical reasoning and thinking rather than algebraic precision. The shift away from mechanics and toward understanding is one attempt to decrease students' anxiety levels, with the assumption that reducing the mathematical content and rote memorization of definitions and formulas reduces students' worries about course performance (Onwuegbuzie, DaRos, & Ryan, 1997).

The importance of using manipulatives has long been maintained (NCTM, 2000). Often, the conventional thinking about manipulatives is that they are useful to school-aged children who are not ready to engage in abstract reasoning and thinking. Researchers have studied the effects of manipulatives on learning mathematics at different grade levels and in different countries (Boggan, Harper & Whitmire, 2010; Castro, 2006; Kelly, 2006). There has been considerable research completed on the use of manipulatives towards the goal of aiding students to better understand mathematical concepts (Bjorklund, 2014; Burns & Hamm, 2011; DeLoache, Scudder & Uttal, 1997; Driscoll, 1983; Freer, 2006; Moyer & Jones, 2006; Raphael & Wahlstrom, 1989; Sowell, 1989; Suydam & Higgins, 1977; Swan & Marshall, 2010). The use of manipulatives in teaching mathematics has developed over time. Golfashani (2013) noted that teaching mathematics has moved away from using beans or counters to using linking cubes, fractions circles and other technologies. Johnson (1993) stated,

> With the increased use of manipulatives, a new attitude is evolving towards mathematics. Mathematics is no longer a set of concrete rules to follow but rather a way of thinking. There are now reasons behind the rules. (p. 11)

Both virtual and concrete manipulatives provide a compelling and promising tool for teaching and learning statistics. The existing literature on virtual and concrete manipulatives applied in education has effectively pointed out the many benefits that they may hold, while recognizing that their effectiveness is primarily reliant on instructor and instructional design. As such, further research on the effects of virtual and concrete manipulatives should focus on these two areas.

Both types of manipulatives are meaningful for learning only with respect to learners' activities and thinking. Physical and virtual manipulatives can be useful, but they will be more so when used in comprehensive, well-planned, instructional settings. Based on Martin (2009), their physicality is not important—their manipulability and meaningfulness make them educationally effective. In addition, some studies suggest that computer manipulatives can encourage students to make their knowledge explicit, which helps them build integrated-concrete knowledge. Such research, using randomized control trials, must be conducted to investigate the specific contributions of concrete and virtual manipulatives to particular aspects of statistics teaching and learning.

Research shows that use of manipulatives over the long term provides more benefits than short-term use does (Sowell, 1989). With long-term use of manipulatives in mathematics, educators have found that students make gains in the following general areas: verbalizing and discussing mathematical ideas and concepts, working collaboratively, thinking divergently to find a variety of ways to solve problems, expressing problems and solutions using a variety of mathematical symbols, making presentations, taking ownership of their learning experiences, and gaining confidence in their abilities to find solutions to mathematical problems (Sebesta &Martin, 2004). It would be beneficial to see if similar relationships apply and hold for the use of manipulatives in other subject areas, especially in secondary education.

All students have different needs in order to maximize their learning, and one type of manipulative can be just as effective as another. The topic, time-frame, type of student being educated, and objective of the lesson all can be factors that play a role in student learning (Tomlinson, 2001). Perhaps combining multiple methods of instruction within a lesson topic could reach more students and make instruction more effective.

Traditional Manipulatives

Definition. Physical (traditional) manipulatives refer to a set of concrete materials that can be physically manipulated by hand. This sensory nature ostensibly makes manipulatives “real,” connected with one’s intuitively meaningful personal self, and therefore helpful (Sarama & Clemets, 2009).

Advantages. Confucius (551–479 BC) once said “I hear, and I forget. I see and I remember. I do and I understand.” Concrete manipulatives could be used to assist students in understanding complex topics. Students having difficulty working on challenging problem-solving tasks may have success when given concrete manipulatives to aide them with the challenge (Jones, 2003). Instructors may find some of the advantages that concrete manipulatives provide easier to apply in their classes. Some of those advantages include: They are more moveable; tactile experience adds a dimension of learning; the student has more control; the process is traceable; depending on the learner type, for some is easier to relate to real-world applications; in some cases, it could be less expensive than technology; students can be more creative; and it allows information to be received visually and kinesthetically (Carbonneau, Marley, & Selig, 2013).

Most of the prior research conducted in this area focused on K-8 classrooms. Phyllis (2001) compared computer and concrete manipulatives for teaching probability concepts to elementary school students and found mixed results. Teachers involved in that study were not convinced that either should be used at the full expense of the other. Klahret al (2006) did a similar comparison in middle school classrooms on an engineering design project and found no differences in learning assessments based on type of manipulative used.

Research conducted by Moyer, 2001 indicated that teachers play an important role in creating mathematical environments that provide students with representatives that enhance their thinking. Vinson (2001) stated, “Using appropriate and concrete instruction materials is necessary to ensure that children understand mathematical concepts” (p. 91).

Swan and Marshall (2010) revisited research on the use of manipulatives in schools. They looked at different ways in which teaching of mathematics and the subsequent learning via the use of manipulatives occurred. Swan and Marshall found that there are potential gains to be made by using mathematics manipulative materials where appropriate and employed in a systematic manner.

Unfortunately, rarely are manipulatives used at the college level, according to the research literature. “Manipulatives help students learn by allowing them to move from concrete experiences to abstract reasoning” and “The effective use of manipulatives can help students connect ideas and integrate their knowledge so that they gain a deep understanding of mathematical concepts” (Boggan, Harper, & Whitmire, 2010, p.4). Manipulatives enhance the abilities of students at all levels to statistically reason and communicate immediately, and the valuable time spent on manipulatives also has sustained, long-term effects on building students’ confidence in learning statistics and deepening their statistics understanding (Shaw, 2002). Working with manipulatives makes practice on skills meaningful and leads to retention and application of information in new problem-solving situations (Klahr, Triona, Williams, 2006). Overall, the indication is that mathematics achievement increases when manipulatives are put to good use.

Through the review of the research on the effectiveness of using manipulatives in teaching undergraduate statistics courses, the authors uncovered numerous studies that supported the use of concrete and virtual manipulatives. For decades, researchers have been demonstrating the positive effects of using concrete manipulatives with their students. Studies that are more contemporary have extended these findings to virtual manipulatives (Hunt, Nipper, & Nash, 2011). As time passes, more and more educational materials (i.e. textbooks, homework assignments, and tests) are available virtually in a digital format. Therefore, it appears to be critical that teachers and instructors of all levels receive the necessary pedagogical training on how to use these materials. Future research on best practices for training teachers using manipulatives would be helpful.

The foundational theory of the study is that when students can visualize a statistical concept in action, a deeper level of understanding occurs. Allen (2007) stated that retention in learning, defined as the ability to retain facts in memory, proves measurable when students have the opportunity to visualize concepts. By giving students concrete ways to view statistics, they can develop relationships between background knowledge and new knowledge (Goracke, 2009).

Disadvantages. Children often can look very busy (active) with manipulatives, but that does not necessarily mean that children are learning. Clements (2000) noted that simply using manipulatives as part of a mathematics lesson does not guarantee success. The results of the present study confirm that for statistics manipulatives to be effective, they must be part of a carefully planned statistics program. The effects of manipulatives upon retention may be investigated by longer-term studies.

Finally, money for purchasing concrete manipulatives could be the number one impediment to the use of virtual manipulatives. Computers may be found in every primary and secondary school and require replacement every three to five years, and yet there seems little concern about the money required to purchase them, maintain them, load software and connect them to printers and the Internet. It is possible that computers are viewed in a different way to manipulative materials and therefore treated differently (Jones, 2003).

Visual Manipulatives

Definition. Virtual manipulatives refer to a set of imagines that can be electronically manipulated on a computer screen. Virtual manipulatives are online versions of physical manipulatives. Moyer, Bolyard, and Spikell (2002) specifically defined virtual manipulatives as "an interactive, web-based visual representation of a dynamic object that presents opportunities for constructing mathematical knowledge" (p.373). The key elements of this definition are that the virtual manipulatives must be web-based, and that it must be manipulatable by the user.

In 2016, Moyer-Packenham and Bolyard revised the definition of virtual manipulatives owing to the rise of technology tools containing virtual manipulatives. Hence, the updated definition of a virtual manipulative is "an interactive technology-enabled visual representation of a dynamic mathematical object, including all of the programmable features that allow it to be manipulated, that presents opportunities for constructing mathematical knowledge." This revision implies that "a virtual manipulative may: (a) appear in many different technology-enabled environments; (b) be created in any programming language; and (c) be delivered by any technology-enabled device" (Moyer-Packenham and Bolyard, 2016, section 1.8).

Advantages. Technology, in the form of virtual manipulatives, in conjunction with the concrete manipulatives, acts as an essential component of enhancing statistics instruction by ensuring students' understanding of statistical concepts. Based on Jones (2003), virtual manipulatives overcome some of the limitations of concrete manipulatives, such as limited materials, but they also come with their own set of challenges.

With the development of Internet, electronic manipulatives have begun to emerge as a teaching and learning aid. Klahr et al. (2006) considered several advantages of electronic manipulatives, including easy access, availability for all, and instant feedback. Virtual manipulatives are one form of electronic manipulatives. Many authors have documented the perceived benefits of virtual manipulatives. A key aspect of these benefits is their availability online and ease of access and management (Dorward 2002; Heath 2002; Leathrum 2001; Moyer & Bolyard 2002). Other benefits of virtual manipulatives are that a large number of developers are able to create and disseminate them and that applets can have a strong focus on specific concepts (Leathrum 2001). Furthermore, virtual manipulatives are capable of doing things that simply are not possible with physical manipulatives, pencil and paper, or other tools (Crawford & Brown 2003; Forster 2006; Keller, Wasburn-Moses & Hart 2002; Reimer & Moyer 2005)

From an instructional standpoint, virtual manipulatives provide students with instant, corrective feedback (Crawford & Brown 2003; Durmus & Karakirik 2006; Reimer & Moyer 2005; Suh & Moyer 2005). Many authors have contended that this ability makes virtual manipulatives well suited to inquiry-based learning and problem solving (Durmus & Karakirik 2006; Jacobs 2005). For example, in their study of fifth graders using a fraction applet, Suh & Moyer found that, "…the applets allowed students to experiment and test hypotheses in a safe environment. The guided format features of the applets allowed guessing and trial-and-error, and at the same time, would not accept and incorrect response" (p. 10).

In addition, virtual manipulatives have the ability to provide multiple representations of a single concept at the same time (Suh & Moyer 2005). Reimer & Moyer (2005) argued that this ability provides an advantage over physical manipulatives. Additionally, it has been proposed that this ability can promote transfer of knowledge from specific ideas to general knowledge (Durmus & Karakirik 2006; Jacobs 2005; Moyer & Bolyard 2002; Suh & Moyer 2005).

Researchers have suggested that use of virtual manipulatives may be helpful for students with disabilities, as well. (Miller, Brown, & Robinson 2002; Riley, Beard, & Strain 2004). In addition, several authors have contended that virtual manipulatives increase motivation and attention in students as well as teachers (Clements & McMillen 1996; Reimer & Moyer 2005; Leathrum 2001).

Very little formal research has been conducted on the effectiveness of virtual manipulatives. Of the research studies addressing virtual manipulatives found for this review, three of them were classroom studies in which two showed some evidence of benefit from using virtual manipulatives, and one showed no difference in using them as opposed to physical manipulatives or no manipulatives at all. No studies address the use of manipulatives in secondary education.

Dorward (2002) conducted a study on the effectiveness of virtual manipulatives in which three groups of students were taught the same topics from three different teachers. One group was taught with physical manipulatives, one with virtual manipulatives and one with no manipulatives. Results on a unit test did not show any differences in student achievement between groups.

Reimer and Moyer (2005) studied a small group of third-grade students learning about fractions with the use of virtual manipulatives. They concluded that virtual manipulatives helped students to learn more about fractions by providing immediate and specific feedback, they were easier and faster to use than paper-and-pencil methods and enhanced students' enjoyment while learning mathematics (Reimer & Moyer 2005, p. 5-6). However, the authors do admit that the small class size and specific demographics fail to make the findings applicable to a broader population.

Suh and Moyer (2005) conducted a similar study of fifth grade students using virtual manipulatives in the classroom for learning about fractions. The authors concluded that virtual manipulatives supported student learning in three important areas: discovery learning, making conjectures, and encouraging students to see mathematical relationships. Again, the specific size and demographics of the class prevent any conclusions from being applied to a larger population.

Keller, Wasburn-Moses, & Hart (2006) studied the use of Java applets for visualizations of 3-D objects in middle and secondary education. In their study, they looked at the effects of use on both students and teachers. They concluded that use of the applets improve students' spatial visualization skills and enhance future teachers' pedagogical content knowledge. Their study highlights an important theme in the literature on physical and virtual manipulatives: teachers play a significant role in the effectiveness of virtual manipulatives.

In the end, any conclusions to be drawn from the above studies are not capable of justifying the use of virtual manipulatives. This fact prompted Reimer and Moyer (2005) to conclude that, "The amount of research on high-quality virtual manipulatives is so limited that a judgment about their potential uses in mathematics instruction is entirely speculative" (p. 8). However, several authors have attempted to justify the use of virtual manipulatives without the use of original research. For example, based on Young (2007), physical manipulatives have been considered effective teaching tools for some time and are supported by a strong research base. The author suggests that this forces consideration of whether the research base supporting physical manipulatives can be directly transferred to the support of virtual manipulatives.

Disadvantages. Clements and McMillen (1996) cited the work of Piaget and Holt to argue that virtual manipulatives are no less concrete than physical manipulatives because both are simply symbolic representations of abstract concepts. Specifically, they argue that the power of manipulatives lies in their concrete nature, and anything that can concretely show an abstract concept helps learning. The logical consequence of this assertion is that the research base supporting physical manipulatives transfers to support the use of computer-based manipulatives (Young, 2007).

In his defense of the validity of the virtual manipulatives found at ExploreLearning.com, Cholmsky (2003) also asserts that Marzano's (1998) meta-analysis of instructional methods that work supports the use of virtual manipulatives. In their previously mentioned study of fifth graders using virtual manipulatives to learn about fractions, Reimer and Moyer (2005) cite the same study from Marzano, specifically regarding graphical/non-linguistic representations, to claim that virtual manipulatives can be an effective learning tool.

Some disadvantages of the virtual manipulatives are: one cannot actually touch them, sometimes forces you to think abstractly, may limit the instructor's ability to follow the students' thought the processes of learning (Sarama & Clements, 2009).

Data on the most commonly used manipulatives will assist postsecondary educators when planning statistics education courses. As with most research, mine raised further questions that require in-depth research. The results implied that one type of manipulative was not better than the other in terms of teaching students within various statistical performance levels (for example: low-achieving, average-achieving, high-achieving).

Methods

The Experiment

Overall, this study is a data analysis of a controlled experiment that was conducted by Dr. William Rayens a few years ago at the University of Kentucky. This experiment included students enrolled in STA 200 sections 022 to 025 in the fall of 2009. STA 200 was a course required of all students who did not take calculus. The class was set up as two large lectures, comprising four 24-person sections, meeting three times a week. Students' ages in STA 200 typically range from 17 to 50, with the majority of the group between 20 and 25 years of age. Gender and ethnic distributions as well as health status are commensurate with the undergraduate population at the University of Kentucky.

For the experiment, two of four sections (with 48 students) were randomly selected for the experimental group and the other two sections (with 48 students) were treated as the control (or business as usual) group, resulting in a total sample size of 96 students. A single calendar of events was created for all four sections. All students had the same lecture. The recitations all followed the same calendar but differed only in the type of manipulative used.

Experiment (EXP) Group. The Experiment Group consisted of 48 students. They had the same lecture as the Business-as-Usual Group. For the recitations, students in the experimental group used virtual manipulatives (i.e., a set of imagines that can be electronically manipulated on a computer screen).

Business-as-Usual Group. The Business-as-Usual Group consisted of 48 students. They had the same lecture as the Experimental Group. For the recitations, students in the Business- as-Usual group used physical (traditional) manipulatives that are concrete and can be physically manipulated by hands.

For example, when students studied the issue of patterned repeated sampling, students in the physical condition spun hand spinners and stacked pegs to create histograms, while students in the virtual condition "spun" virtual spinners on computer screens and stacked virtual pegs to create their histograms. Essentially, the two types of manipulatives were used as helping-to-learn tools for key conceptual constructs such as Central Limit Theorem (through Spinning Bells activity), Experimental Design (through Whacking Moles activity), Probability and Area (through Corn Hole Likelihood activity) and Confidence Intervals (through Confidence in Repetition activity). These activities were well constructed and available from Dr. William Rayens at the University of Kentucky (rayens@uky.edu).

The Data

The experiment went on for the entire semester. Data on various variables was collected throughout the semester. A cognitive assessment on the identified conceptual constructs was developed, introduced in class, pilot tested, and revised. An affective assessment on the level of student engagement in learning (i.e., the course) was also developed. Both assessments were given to students in both conditions. These instruments were used to test a null hypothesis that there are no differences in these cognitive and affective measures with respect to different manipulatives used. As part of this cognitive assessment, the same final test was administered to students in both groups, in the paper-and-pencil format including multiple-choice and short-answer items (120 minutes of testing time for 49 items). During the semester, midterm test scores and two-minute assignment scores were also collected.

To enrich the data that has already been collected for more fruitful analysis, variables related to student background were incorporated into the existing data. Student background variables included gender, race, SAT mathematics scores, and cumulative GPA for the first and second years at the University of Kentucky. These variables functioned mainly as control variables in data analysis.

The Analysis

Multiple correlation/regression analysis was used to test the between-group differences in cognitive and affective learning outcomes between students in virtual and physical conditions. Statistical analysis contained two related components. To examine the short-term effects (i.e., at the end of the semester when the experiment was implemented), a multiple regression approach to ANOVA was adopted. Statistics achievement was the dependent variable. It is the course average, a combined measure with equal weights of the final and midterm tests as well as the two-minute assessments on the key topics of statistical vocabulary, confidence intervals, hypothesis testing, experimental design, sampling distributions, generic normal calculations and correlation. Student characteristics included continuous variables of age and high school mathematics ACT score as well as dichotomous variables of gender. All student variables were self-explanatory in meaning. Three preliminary models tested main effects of types of manipulatives, main effects of students' gender, age, and high school ACT mathematics scores (linear and quadratic terms) respectively, and the interaction effects between types of manipulatives and students' gender, age, and high school ACT mathematics scores (linear and quadratic terms) respectively.

To examine the long-term effects of the treatment (i.e., one year after completing the course), multiple regression approach to ANOVA was adopted. Students' grade point average (GPA) one year later was the dependent variable with two key independent variables. The first one was the type of manipulatives used in teaching the class (traditional concrete versus online virtual). The second one was the course average score for that semester. In particular, the interest was in testing the effects of type of manipulatives and course average scores on GPA one year later, as well as the interaction between type of manipulatives and course average scores.

RESULTS

As stated earlier, this dissertation uses a randomized experiment to determine if differences in students' achievement in undergraduate statistics class exist when students learn statistical concepts using virtual manipulatives compared to when students learn statistical concepts using physical manipulatives. The researcher randomly assigned students in different sections to either a physical manipulative condition or a virtual manipulative condition. For the experiment, two of four sections (with 48 students) were randomly selected for the experimental group, and the other two sections (with 48 students) were treated as the control (or business-as-usual) group, resulting in a total sample size of 96 students. A single calendar of events was created for all four sections. All students had the same lecture. The recitations all followed the same calendar but differed only in the type of manipulative used.

Table 1 presents descriptive statistics of the two outcome variables, course average score and second-year grade point average, as well as student characteristics including gender, age and high school ACT mathematics score. The table shows 2.16 points difference in course average score in favor of the Business-as-Usual group (BAU) in comparison with the Experimental group (EXP), and no difference in second-year grade point average between the EXP and BAU group. There are 21% males (SD=0.06) in the BAU group and 33% males (SD=0.07) in the EXP group. Age distribution for both groups was very similar, with M=24.15, SD=0.23 for the BAU group versus M=24.02, SD=0.26 for the EXP group. The high school ACT mathematics score distribution for both groups is also very similar with M=22.24, SD=0.67 for the BAU group versus M=23.41, SD=0.71 for the EXP group.

With Tables 2, 3 and 4, we try to predict course average score from variables by means of a regression analysis. Class performance is our dependent variable. The key independent variable is the treatment dummy comparing EXP with BAU. Gender, age and ACT mathematics score are control (independent) variables respectively. The coefficients (effects) tell us how many units the course average score increases for a single unit increase in each independent variable. Three preliminary models tested main effects of treatment (types of manipulatives) and students' gender, age, and high school ACT mathematics scores, as well as their interaction effects on course average score.

Table 2 presents results of preliminary model testing treatment effects between online virtual manipulatives and traditional concrete manipulatives with consideration of student gender. The course average score for a male student who used traditional manipulatives was 85.16. This average is statistically significantly different from zero. The interaction effects between virtual manipulatives and gender were not statistically significant. There were not statistically significant treatment main effects; neither were there statistically significant gender main effects. The regression model overall accounted for 6% of the variance in course average score.

Table 3 presents results of preliminary model testing treatment effects between online virtual manipulatives and traditional concrete manipulatives with consideration of student age. Student age was centered around its grand mean. Therefore, the course average score for a student of average age who used traditional manipulatives was 84.29. This average is statistically significantly different from zero. The interaction effects between virtual manipulatives and student age were not statistically significant. There were not statistically significant treatment main effects; neither were there statistically significant age main effects. The regression model overall accounted for 5% of the variance in course average score.

Table 4 presents results of preliminary model testing treatment effects between online virtual manipulatives and traditional concrete manipulatives with consideration of high school

ACT mathematics score. The high school ACT mathematics score was centered around its grand mean. Therefore, the course average score for a student with an average high school ACT mathematics score who used traditional manipulatives was 85.57. This average is statistically significantly different from zero. The interaction effects between virtual manipulatives and high school ACT mathematics scores were not statistically significant. There were not statistically significant treatment main effects; neither were there statistically significant high school ACT mathematics score main effects. The regression model overall accounted for 10% of the variance in course average score.

In sum, in each model, there are neither statistically significant interaction effects nor statistically significant treatment effects (as main effects of types of manipulatives).

Table 5 presents the simplified results for the treatment effects of online virtual manipulatives against traditional concrete manipulatives in terms of course average score. This table aimed to examine the short-term effects of the treatment. The course average score for a student who used traditional manipulatives was 84.22. The model showed a statistically insignificant treatment effect, which indicates that statistics achievement measured with course average scores in the traditional concrete manipulative group was statistically no different from that in the online virtual manipulatives group. The regression model overall accounted for 1% of the variance in course average score.

For Table 6, the interest was in the long-term effects of types of manipulatives and course average score on a grade point average one year later, measured by testing the effects of type of manipulatives and course average scores on Grade Point Average (GPA) one year later, and the interaction effects between type of manipulatives and course average scores. The GPA for a student who used traditional manipulatives and was average achieving in the course was 3.13 one year later. This average is statistically significantly different from zero. There is statistically insignificant interaction effect between the types of manipulatives used and course average score on GPA one year later. There were statistically insignificant differences in GPA one year later between the students using concrete traditional manipulatives versus virtual online manipulatives. However, the course average scores have a statistically significant effect on GPA one year later. A one-point increase in the course average score is associated with an increase of 0.05 points in GPA one year later, holding the rest of the predictors constant. The regression model overall accounted for 56% of the variance in GPA one year later.

Table 7 presents the results for effects of types of manipulatives and course average score on grade point average one year later, with control for student characteristics. The GPA for a female student of average age with average high school ACT mathematics score who used traditional manipulatives and was average achieving in the course was 3.31 one year later. This average is statistically significantly different from zero. After control for student characteristics, there is statistically insignificant interaction effect between the types of manipulatives used and the course average score on GPA one year later. There were statistically insignificant differences in GPA one year later between students using concrete traditional manipulatives versus virtual online manipulatives, after control for student characteristics. However, after control for student characteristics, the course average scores have a statistically significant effect on GPA one year later. A one-point increase in the course average score is associated with an increase of 0.03 points in GPA one year later.

In addition, there are two statistically significant predictor variables concerning student characteristics in this model (gender and high school ACT mathematics score). The GPA of females is 0.27 points higher than that of males one year later (while holding all other predictors in the model constant). One-point increase in high school ACT mathematics is associated with an increase of 0.03 points in GPA one year later (while holding all other predictors in the model constant). The regression model overall accounted for 69% of the variance in GPA one year later.

DISCUSSION

Summary of Principal Findings

The results of this study revealed that there were no significant differences between the BAU group who received traditional concrete manipulatives and the EXP group who received online virtual manipulatives. This study included high school ACT mathematics scores. Nonetheless, there were no statistically significant interaction effects between types of manipulatives and ACT scores. This study informs the literature that ability levels neither intensify nor weaken the effects of types of manipulatives.

The result of no significant difference in GPA one year later refers to the exploration of the long-term effects of types of manipulatives and performance in that course. The results of the study did not show a significant difference in GPA one year later between the EXP group and the BAU group. The results of the study did demonstrate, nonetheless, that performance (regardless of types of manipulatives) in that course had a positive impact on GPA one year later.

Insights to Research Literature

The foundational position of the study is that when students can visualize a statistical concept in action, a deeper level of understanding occurs. By giving students concrete ways to view statistics, they can develop relationships between background knowledge and new knowledge (Goracke, 2009). We believe that technology, in the form of virtual manipulatives, perhaps in conjunction with the concrete manipulatives, acts as an essential component of enhancing statistics instruction by ensuring students' understanding of statistical concepts. This study thus compared the effectiveness of using concrete and virtual manipulatives in undergraduate-level statistics class. Virtual manipulatives and traditional manipulatives are equally effective and do not produce long-term differences academically. These results illustrate some advantages of the use of virtual manipulatives, such as the ability to provide feedback to students immediately upon rendering their response (i.e., instant feedback). Feedback must be administered in a timely fashion in order to lend value to the learning environment (Crompton, 2011). After receiving immediate feedback, students can rethink their course of action and collaborate with classmates on an alternative process to reach a solution. Virtual manipulatives are also dynamic, interactive, flexible and easy to manage. They make an interesting complement to concrete manipulatives. The advantages of their use in the classroom are promising in the search for new ways of teaching and learning statistics.

When it comes to manipulatives, their physicality is not important—their manipulability and meaningfulness make them educationally effective supports (Martin, 2009). On this point, this study offers more support. In addition, some studies suggest that computer manipulatives can encourage students to make their knowledge explicit, which helps them build integrated-concrete knowledge, but rigorous causal studies have not been conducted to our knowledge (Sarama & Clements, 2016).

Implications for Educational Policies and Practices

The results of this study support Clements' (1999) hypothesis that computers can provide students with virtual representations of statistical concepts that are just as meaningful as physical manipulatives. Specifically, Clements hypothesized no difference between virtual and physical representations in the mathematics classroom. They also align with the results of previous empirical science studies that compare virtual and physical manipulatives (Klahr, Triona, & Williams, 2007; Moyer, Niezgoda, & Stanley, 2005; Reimer & Moyer, 2005; Smith, 2006; Steen, Brooks, & Lyon, 2006; Suh, 2005; Suh & Moyer, 2007; Triona & Klahr, 2003). Overall, when facing a shortage of traditional manipulatives (e.g., due to funding), instructors may take advantage of the easy access to virtual manipulatives through the internet.

Because virtual and physical manipulatives are equally effective, an educational issue to consider is how to increase teachers' awareness of and abilities to effectively use virtual manipulatives for teaching (Crawford & Brown 2003; Gadanidis, Gadanidis, Schindler 2003; Reimer & Moyer 2005). For example, Moyer (2001) argued that taking an interest in virtual manipulatives is not enough without considering how it is used in the classroom. Virtual manipulatives should always be created in conjunction with a study of their classroom application in order to provide tools that take into account the students' needs.

Limitation and Suggestion for Further Research

The purpose of this study was to compare the effectiveness of concrete manipulatives and virtual manipulatives when teaching statistics to undergraduate students in a core content course. The results of the study may not be generalized to the population of all undergraduate students. It also gives no indication of how the results would generalize to other content domains. Further studies may explore along these lines of inquiry regarding the effects of virtual manipulatives in comparison with concrete manipulatives.

Research shows that consistent use of manipulatives provides more benefits than temporary use (Sowell, 1989). With consistent use of manipulatives in mathematics, educators have found that students make gains in the following general areas (Sebesta and Martin, 2004): verbalizing mathematical thinking; discussing mathematical ideas and concepts; relating real-world situations to mathematical symbolism, working collaboratively, thinking divergently to find a variety of ways to solve problems, expressing problems and solutions using a variety of mathematical symbols, making presentations, taking ownership of their learning experiences, and gaining confidence in their abilities to find solutions to mathematical problems using methods that they come up with themselves without relying on directions from the instructor. With only one semester's use of virtual manipulatives measured in this study, the duration may not be classified as consistent use. Further studies may seek some longer period of using virtual manipulatives.

Table 2.1

Descriptive Statistics of Outcome Variables and Student Characteristics

	BAU (n = 48)		EXP (n = 46)	
Variable	Mean	SD	Mean	SD
Course average score (continuous)	84.44	1.08	82.28	1.74
Second-year grade point average (GPA) (continuous)	3.17	.07	3.17	.09
Male (= 1 versus female = 0)	.21	.06	.33	.07
Age (continuous)	24.15	.23	24.04	.26
High school ACT mathematics score (continuous)	22.24	.67	23.41	.71

Table 2.2

Results of Preliminary Model Testing Treatment Effects between Online Virtual Manipulatives and Traditional Concrete Manipulatives with Consideration of Gender

	Effects	SE
Constant	85.16*	1.57
A: Online virtual (vs. traditional)	-1.03	2.34
B: Gender	-4.50	3.44
A × B	-1.17	4.59
Proportion of variance explained	0.06	

* $p < .05$.

Table 2.3

Results of Preliminary Model Testing Treatment Effects between Online Virtual Manipulatives and Traditional Concrete Manipulatives with Consideration of Student Age

	Effects	SE
Constant	84.29*	1.40
A: Online virtual (vs. traditional)	-2.07	2.01
B: Age	-1.40	.87
A × B	.36	1.19
Proportion of variance explained	0.05	

* $p < .05$.

Table2. 4

Results of Preliminary Model Testing Treatment Effects between Online Virtual Manipulatives and Traditional Concrete Manipulatives with Consideration of High School ACT Mathematics Scores.

	Effects	SE
Constant	85.57*	1.73
A: Online virtual (vs. traditional)	-3.59	2.31
B: High school ACT mathematics	.40	.45
A × B	.39	.56
Proportion of variance explained	0.10	

* $p < .05$.

Table 2.5

Results of Simplified Model Estimating Treatment Effects between Online Virtual Manipulatives and Traditional Concrete Manipulatives (in Terms of Course Average Scores)

	Effects	SE
Constant	84.22*	1.42
Online virtual (vs. traditional concrete)	-1.95	2.03
Proportion of variance explained	.01	

* $p < .05$.

Table 2.6

Effects of Types of Manipulatives and Course Average Scores on Grade Point Average (GPA) One Year Later without Control of Student Characteristics

	Effects	SE
Constant	3.13*	.05
A: Online virtual (vs. traditional concrete)	.08	.08
B: Course average scores	.05*	.01
A × B	-.01	.01
Proportion of variance explained	.56	

* $p < .05$.

Table 2.7

Effects of Types of Manipulatives and Course Average Scores on Grade Point Average (GPA) One Year Later with Control for Students Characteristics

	Effects	SE
Constant	3.31*	.06
A: Online virtual (vs. traditional concrete)	-.08	.07
B: Course average scores	.03*	.01
A × B	.01	.01
Gender	-.27*	.09
Age	.01	.04
High School ACT mathematics	.03*	.01
Proportion of variance explained	.69	

* $p < .05$.

CHAPTER 3

THE EFFECTS OF AN INVERTED INTRODUCTORY STATISTICS CLASSROOM ON LEARNING OUTCOMES OF UNDERGRADUATE STUDENTS

Despite growing efforts in improving the teaching and learning of statistics at all educational levels, few direct connections have been established between research and practice (Garfield & Zvi, 2009). Educational research has long been interested in the assessment of statistics anxiety and attitudes toward statistics as well as other factors (e.g., mathematical background and motivation to learn) that predict student achievement in statistics (see Garfield & Ben-Zvi, 2007). However, only recently have researchers started to investigate the understanding and reasoning of students concerning critical statistical concepts, particularly how these concepts can be developed through a carefully planned sequence of learning activities and how this strategy can be implemented effectively in the classroom setting (Garfield & Zvi, 2008). Obviously to address this issue, empirical studies, particularly experiments in real educational settings such as a university classroom, are needed.

The main goal of this study is to fill in some gaps in the research literature on the teaching and learning of statistics at the college level. This study joins the reform effort that seeks alternative content and pedagogy in statistics education. To promote the link between research and practice, educational experiment will be used in this study to examine the effects on learning outcomes of different instructional practices in statistics education, in particular the use of different styles of instruction. Specifically, this study will investigate the impact of different teaching methods for presenting statistical content (i.e., inverted classroom versus traditional classroom) on learning outcomes in introductory statistics. Some student background factors (e.g., prior ability) and course structure factors (e.g., availability of extra credits) will also be brought into the equation to examine whether they are capable of enhancing treatment (intervention) effects.

This study attempts to test the conditional effectiveness of the emerging teaching method of inverted classroom in statistics education. In other words, this study seeks to produce inferential and descriptive statistical evidence to assess the effectiveness of the inverted classroom. The following research questions will be addressed:

1. Are there any differences between inverted classroom and traditional classroom in terms of cognitive (performance in statistics) and affective (attitude toward statistics) outcomes of undergraduate students in introductory statistics?
2. Are there any important student background variables and course structure variables that are able to enhance treatment (intervention) effects (associated with inverted classroom) on cognitive and affective outcomes of undergraduate students in introductory statistics?

Overall, with the increased popularity of the inverted classroom, the results from this controlled experiment will add substantially to our limited knowledge base regarding this instructional practice. The results of this study will improve undergraduate statistical education by discovering meaningful links between research and practice. This study joins many other reform efforts to explore instructional ways that engage students in reasoning and thinking statistically.

Literature Review

The Call for Inverted Instruction

Inverted learning is an instruction design that replaces the traditional model of lecturing in class and assigning practice problems for homework with a model of assigned learning activities for homework and practice problems in class, hence the term inverted classroom (Hamden, McKnight, McKnight, & Arfstrom, 2013). Other terms have been used for this design, such as interactive teaching (Walvoord & Anderson, 1998) and inverted teaching (Lage et al., 2000; Strayer, 2012). Blended, hybrid, and e-learning are other terms circulating in literature that share some similarities with flipped classrooms, but refer to the mixing of face-to-face class time with online learning (Snart, 2010). In blended or hybrid classes, there is a trade-off of class time with online learning components such as discussion boards. In inverted classrooms, there is no trade-off; class time is still preserved as a whole-group meeting, albeit students may work in peer groups within the whole-group session.

Goodwin and Miller (2013) have commented that evidence on flipped classrooms is being reported on social media websites, but not at a sufficient level to be called research-based. For the inverted classroom to become an accepted instructional practice, more scientific research should be conducted on its effects, positive or negative. This research addresses these deficiencies in current inverted classroom research by investigating an inverted classroom for undergraduate college statistics class. The detailed description of the design, implementation, and evaluation of an inverted undergraduate college statistics classroom contained in this document may benefit instructors, departments, institutions, agencies, and governments.

In higher education, problems with undergraduate college statistics are especially alarming. College statistics is a required course for diverse majors, but it is viewed by many as a gatekeeper course, controlling student access to degree completion (Reyes, 2010). College-level statistics provides the foundational skills, conceptual understanding, and mathematical insights needed for success in subsequent courses (Dugopolski, 2010). Application problems and investigations of modeling allow students to see how useful college statistics is in the real world.

Online searching that involved several keyword combinations and research databases failed to identify any literature regarding the use of an inverted classroom in large undergraduate statistics course. The complete lack of research literature regarding the application of this teaching method in college statistics classrooms implies that much remains to be learned, making the inverted classroom for college statistics a productive research topic. Moreover, the emergence of the inverted classroom introduces a possibility of significantly improving undergraduate statistics education. This possibility, combined with the evident absence of research literature concerning the use of the inverted classroom in college statistics, has created a pressing need for available information regarding this topic.

The Role of Inverted Instructor

The flipped classroom involves a very important transformation of the teacher's role. In a traditional class, the teacher can be described as the "sage on the stage" who presents information in engaging ways in hopes that students will pay attention and absorb the information (Bergmann, Overmyer, & Wilie, 2012). The flipped classroom moves away from this idea, placing the teacher in the role of the "guide on the side" who works with the students to guide them through their individual learning experiences (Bergmann, Overmyer, & Wilie, 2012). The "guide" role can be illustrated using Paulo Freire's idea that education "should not involve one person acting on another, but rather people working with each other," (Smith, 2012, p. 1).

The flipped classroom requires that the instructor create an inquiry-based teaching environment, where the face-to-face class time shifts from a teacher-centered space to a student-centered space (Bergmann & Sams, 2012a). The traditional educational system was created using the factory model of management with the idea of top-down instruction, and "sage on the stage" teachers who produce outputs, or students who pass standardized tests (Howell, 2013). However, a paradigm shift is occurring where learning begins to be about students and their needs. “Since the turn of the century, the challenges of globalization, information technology, international competition, and strong local developments have stimulated a new wave of educational reforms” (Cheng & Mok, 2008, p. 374). The new wave has shifted from a teacher-centered paradigm to a student-centered one. Cheng and Mok (2008) described this new paradigm as one where learning should be tailored to meet the needs of the individual student. Kirch (2012, p. 4) reports that the flipped classroom ideology has allowed her to “interact with every student (all of them) on a daily basis in at least a short conversation” and “be able to more easily and readily assess student mastery of the content on a daily basis and provide the immediate support they need to succeed.”

One study looks at a flipped model in AP Calculus (Strauss, 2012). The instructor created about four videos per week with a length of 20 to 30 minutes each. This method is unique, in that the videos were not all created far in advance but were often created only a few days before use. This allowed the instructor to customize the videos based on the progress of the course (Roshan, 2011). The following shows the AP Calculus exam results (Roshan, 2011):

> In surveys administered by Byron High School mathematics teachers, 87 percent of parents and 95 percent of students said that they preferred flipped learning to the traditional lecture format for mathematics. Many students commented that they prefer interacting with others during class time, having help available in class, and having the ability to re-watch the lectures if needed. Faulkner also stated that because of the increased one-on-one time with students in class, teachers and students were able to build better relationships. (p. 41, Overmyer, 2014)

The Operation of Inverted Instruction

Inverted classroom refers to the instructional practice where events that traditionally take place inside of the classroom now take place outside of the classroom and vice versa. For example, students in the inverted condition are required to watch video lectures before coming to class. This typically happens in the classroom but now becomes their homework. When students come to class, they may complete activities that are designed to help them engage in discovery learning of the content already experienced by watching the videos. This is usually what students do on their own after class, but now students interact with each other and the instructor in class as they work to deepen their understanding. Obviously, in an inverted classroom, the main class time no longer involves a traditional lecture (Bishop & Verleger, 2013; Lage, Platt, & Treglia, 2000; Strayer, 2012).

Traditional classroom refers to statistics classes that are taught using the traditional lecture method. Typically, students come two times a week to a classroom and listen to a lecture about certain statistical content. Often, these traditional lectures are heavily content driven where the instructor introduces statistical concepts and then works though examples that apply those concepts. During the lectures, students may have opportunities to ask questions and answer questions from the instructor related to the content under discussion. It is possible for lectures in a traditional classroom to be presented in an interactive manner. Obviously, in a traditional classroom, the main class time involves a traditional lecture (Lage, Platt & Treglia, 2000; Strayer, 2012; Wilson, 2013).

The Brief History of Inverted Instruction

The concept of the flipped classroom and flipped learning is not new (Baker, 2000; Strayer, 2007). Before flipped classrooms, distance learning utilized instructional videos to deliver content. The idea that new technologies such as television and radio could be used to deliver education began to surface as long ago as the 1920s (Byrne, 1989). The Open University was the first and most successful effort to use video to broadcast educational content. This practice began in the 1960s in the United Kingdom to address the elimination from higher education of people from lower income groups. Originally, the Open University was the "University of the Air," a daily distance education television program seen in the early morning throughout the United Kingdom, Canada, and Australia (The Open University, 2013).

As early as 1982, Baker had a vision of using electronic means to "cover" rote material outside of class (Baker, 2000). The basic concept he applied in his class was to move the rote transmission of information that had been the content of his lectures out of the classroom and to use the opened-up class time for the students to work on application of the principles from that content while he was there to see what they were doing, answer questions and make suggestions. (Baker, 2011, p. 2) Detailed student comments indicated that the learning was more personalized, the cooperative groups promoted critical thinking, and the online resources gave students more control over their learning. Baker presented the concept to conferences between 1996 and 1998, and in 1998 began to refer to the method as "The Classroom Flip" (Baker, 2011).

At approximately the same time, Lage, Platt and Treglia (2000) designed and implemented a similar method. Lage, Platt, and Treglia (2000) used the phrase "inverted classroom" in their study of the perceptions of students and instructors in introductory economics courses. They referred to the concept as "The Inverted Classroom" and similarly held the expectation that students would view lectures in advance of class, then spend class time clarifying difficult concepts and working in small groups. The use of learning technologies, particularly multimedia, provide new opportunities for students to learn," (p. 32). They used the inverted teaching method on five sections of an economics course. The inverted classroom was presented by Lage et al. in their 2000 seminal work “Inverting the Classroom: A Gateway to Creating an Inclusive Learning Environment.” In their article, Lage et al. included details of the selected instructional strategy and reported on student perceptions and engagement for this early inverted classroom.

The modern use of online videos to supplement face-to-face instruction is often credited to Bergmann and Sams (Pink, 2010). In 2007, they were both science teachers at Woodland Park High School in Colorado. According to Bergmann and Sams (2012a), the early recordings were only for students who missed class:

> Our absent students loved the recorded lectures. Students who missed class were able to learn what they missed. Some students who were in class and heard the live lecture began to re-watch the videos. … And we loved it because we didn’t have to spend hours after school, at lunch, or during our planning time getting kids caught up. (p. 1)

After scouring the Internet, they found that no one else was utilizing this method. The name was briefly changed to reverse instruction, but then, in 2010, Dan Pink wrote about the method and called it the “flipped classroom,” and the term has stuck (Bergmann & Sams, 2012a). Bergmann and Sams have popularized the “flipped learning” pedagogy through the creation of The Flipped Learning Network website, and by writing the book, *Flip Your Classroom: Reach Every Student in Every Class Every Day* (2012). Since 2009, Woodland Park High School has hosted a summer workshop for educators interested in the flipped learning model. Each year, attendance has risen sharply, and in June 2012, flipped educators provided a workshop for more than 500 attendees (Overmyer, 2013).

A flipped classroom is not necessarily a new method of teaching. Rather, it is an older idea that has become more organized and has attracted more attention as educators search for more effective ways to teach. Opinions in the education community regarding the flipped classroom are mixed. Some educators consider the flipped classroom to be the future standard of educational technique (Bergmann, Overmyer, & Wilie, 2012). Other educators consider the flipped classroom to be a passing trend that will be found to be an ineffective and undesirable form of education (Bergmann, Overmyer, & Wilie, 2012).

The Conference Board of Mathematical Sciences (CBMS) 2010 survey (American Mathematical Society [AMS], 2013f) found that, at public two-year colleges, 79 percent of college algebra sections and 89 percent of college algebra and trigonometry (combined) sections were taught mostly by the standard lecture method. Because of the high failure rates of this traditional teaching approach, other pedagogical methods are being explored (Baxter Hastings et al., 2006). More student-centered approaches are being promoted, which encourages more student engagement (Huba & Freed, 2000). As such an alternative, the inverted or flipped classroom model is receiving increased attention in educational circles and popular press (Toppo, 2011; Tucker, 2012). The human interactions that now occur in the classroom are the most significant aspects of the inverted classroom model (Bergmann & Sams, 2012b). Strayer (2007) reported that in most instances where the classroom flip is used, the goal is to create an active learning environment during class meetings, while ensuring content coverage. Strayer's conceptual framework is derived from Piaget's theories of active learning. The classroom flip is usually motivated by a desire to learn through active participation in the classroom. Piaget says that learning occurs not when a person merely copies an idea, but when a person acts on it. (Strayer, 2007, p. 45)

With the growth of Internet technology, virtual communications, and learning management systems, many educators are interested in an inverted classroom (Berrett, 2012). In 2007, science teachers Jonathan Bergman and Aaron Sams created a new movement in education called the flipped classroom method. Recently, articles on this topic have appeared in *USA Today* (Dell Cava, 2012), *The New York Times* (Rosenberg, 2013), *The Economist* (Flipping the Classroom, 2011), and *The Washington Post* (Strauss, 2012). In early 2010, a professional learning network was created for educators interested in the flipped model. As of April, 2016, the network had more than 28,000 members worldwide. This network provides both pedagogical and best-practice discussions as well as pragmatic support on technology and implementation (Overmyer, 2013).

Most of the studies on the flipped classroom model in undergraduate education were in the STEM fields. This is not surprising since these are the subjects that are most commonly flipped (Overmyer, 2013). Strayer (2012) compared the learning environments of a flipped introductory statistics class with a traditional introductory statistics class. Students in the flipped classroom were less satisfied with how the classroom structure oriented them to the learning tasks in the course, but they became more open to cooperative learning and innovative teaching methods. Strayer showed that students in a flipped classroom environment preferred the method and displayed a higher level of innovation (being able to solve problems in creative and unique ways) and cooperation (familiarity with working with others to solve problems and discuss ideas) than students in a traditional classroom setting. His results also indicate that students in a flipped classroom experience a lower level of task orientation than students in a traditional classroom (Strayer, 2008). From the results of his study, Strayer gives recommendations for the implementation of flipped classrooms for undergraduate level courses. One of them is to provide step-by-step instructions for classroom activities to create more structure for the students (Strayer, 2008). To create more structure, a teacher also could scaffold the activities, suggests Strayer. Scaffolding is instruction given when learning a new task where different levels of support are given, with student eventually having most or all support removed as the activity progresses (Hogan & Pressley, 1997). Another recommendation is to keep open activities short, spending no more than two lessons on any one activity (Strayer, 2008). According to Strayer, one effect of the flipped classroom is that students become more aware of their own learning processes (Strayer, 2008). Because of this increased awareness, students will need more time to reflect upon their activities to make connections to the course material (Strayer, 2008).

Another study on the flipped classroom was conducted by Toto and Nguyen. In this flipped classroom, students watched a 30-minute video lecture prior to going to class. As a result, there was additional free time in class, which was spent using real-world tools and engaging in practical applications (Toto & Nguyen, 2009). This classroom was found to have increased student engagement (Toto & Nguyen, 2009). Furthermore, students had more opportunities to gain a sense of how the tools and ideas they were leaning are used in the real world (Toto & Nguyen, 2009).

The Major Advantages of Inverted Instruction

There are three primary motivations for using an inverted classroom based on Mason, Shuman and Cook, 2013. First, the inverted classroom frees class time for interactive activities, such as active, cooperative, and problem-based learning, and for reinforcing course material without sacrificing content (Zappe, Leicht, Messner, Litzinger, Lee, 2009). Second, the inverted classroom allows an educator to present course material in several different formats, and so engage the students' various learning styles and preferences, (Zappe, Leicht, Messner, Litzinger, Lee, 2009), (Lage, Platt, Treglia, 2000). Third, the inverted classroom can encourage students to become self-learners and help prepare them for how they will need to learn as practicing engineers (Bland, 2006).

Kathleen Fulton (2012) listed the following among the advantages of the flipped classroom: (1) students move at their own pace; (2) doing "homework" in class gives teachers better insight into student difficulties and learning styles; (3) teachers can more easily customize and update the curriculum and provide it to students 24/7; (4) classroom time can be used more effectively and creatively; (5) teachers using the method report seeing increased levels of student achievement, interest, and engagement; (6) learning theory supports the new approaches; and (7) the use of technology is flexible and appropriate for “21st century learning;” (8) there is more time to spend with students on authentic research; (9) students get more time working with scientific equipment that is only available in the classroom; (10) students who miss class for debate/sports/etc. can watch the lectures while on the road; (11) the method “promotes thinking inside and outside of the classroom;” (12) students are more actively involved in the learning process; and (13) they also really like it.

Teachers using the flipped method say its primary benefit is that, for the first time in their teaching careers, they have some one-on-one contact with every student during every class period (Moore, Gillett, & Steele, 2014). Ideally, the flipped model is a blending of direct instruction with inquiry-based learning. This allows more time for the development of 21st century skills such critical thinking, collaboration and self-direction (*Framework for 21st Century Learning,* 2010). *The Flipped Manifest* (Bennet, et al., 2011) states that:

> Practitioners of the various flipped classroom models are constantly tweaking, changing, rejecting, adding to, and generally trying to improve the model through direct experience with how effective it is for kids. It's not "record your lecture once" and you're done; it's part of a comprehensive instructional model that includes direct instruction, inquiry, practice, formative and summative assessment and much more. It also allows teachers to reflect on and develop quality and engaging learning opportunities and options for internalization, creation, and application of content rather than just fluff or time filling assignments. (p.1)

The flipped method also may have benefits for at-risk students. One example is an economically challenged school near Detroit with 75 percent of students receiving free or reduced-price lunch, with many students commuting from Detroit. The main issue facing the school was failure and drop-out rates (Pearson Case Study, 2013). The school's principal reversed the instructional procedures so that students did homework at school. The flipped classroom model also may have benefits in reducing anxiety in difficult, content heavy courses. *The Washington Post* (Strauss, 2013) article details Stacy Roshan who teaches AP Calculus using the flipped classroom model at a private school in Potomac. According to Roshan, the traditional classroom for this course is a really anxious environment with too much material and not enough time. The flipped classroom allows her to remove the lecture from the classroom and provide one-on-one time with students in the classroom. Students like the method because they no longer have to sit at home and struggle with confusing homework. In the end, students feel that it is much easier to learn calculus, and that the method has reduced their math anxiety.

Although there is no prescribed method for flipping a classroom because of the abundance of instructional strategies that can occur inside and outside of the classroom, there are, according to Brame (2013), four common key elements indicative of the flipped classroom: an opportunity for students to gain first exposure prior to class' an incentive for students to prepare for class' a mechanism to assess student understanding' and in-class activities that focus on higher level cognitive activities. These elements are the backbone of a flipped classroom, and each one is tied to important learning principles that make the flipped classroom a potential teaching method that can improve student learning. Supporters argue that the videos maximize class time to promote the exact deeper, inquiry-based learning that the critics bemoan (The Economist, 2011). Proponents of the flipped model argue that how a teacher uses the newly freed class time is most important (Bergmann & Sams, 2012a).

Zappe, Leicht, Messner, Litzinger, and Lee (2009) flipped a large undergraduate architectural engineering course. Based on student evaluations of the course, the authors indicated that the classroom flip had a positive impact on student learning. Students perceived the method of teaching as more effective than lecturing and reported that they enjoyed the class and benefited from watching the lecture videos outside of class (Zappe, Leicht, Messner, Litzinger, Lee, 2009).

Ruddick (2012) applied the flipped classroom concept to a college preparatory chemistry course. Results showed that the inverted classroom students outperformed the standard lecture-based students, with higher final exam scores and overall success in the class (Ruddick, 2012). Based on Ruddick, the inverted classroom students became more interested in and felt less intimidated by chemistry and found the online video and PowerPoint materials useful.

Some current studies in specific academic disciplines and levels offer evidence that the flipped classroom model is beneficial in undergraduate education and worthy of future research. Love, Hodge, Grandgenett, and Swift (2013) reported that sophomores in an experimental flipped applied linear algebra course did as well as students in a traditional lecture-based course on common final exams, but students from the flipped class enjoyed class more than those in the lecture-based course. Although student scores were not higher, the researchers gave commendation to the flipped classroom method because it left students with a more positive attitude toward mathematics, an admirable consequence in light of the goal to increase interest in STEM areas in undergraduate education.

The use of the inverted classroom has the potential to be an effective and beneficial method of education. Replacing direct instruction during class time with video lectures observed outside of the classroom allows for more class time to be used for active learning. Active learning can include activities, discussion, student-created content, independent problem solving, inquiry-based learning, and project-based learning (Bergmann, Overmyer, & Wilie, 2012). This use of class time can create a classroom environment that uses collaborative and constructivist learning, blending with direct instruction used outside the classroom (Tucker 2012). Constructivist learning takes place when students gain knowledge through direct personal experiences, such as activities, projects, and discussions. (Ultanir, 2012). The frequency of these personal experiences can be increased in a flipped classroom through the use of activities, creating students who are active learners rather than passive learners (Minhas, Ghosh, & Swanzy, 2012; Sams, 2013). The passive learning of a flipped classroom happens during the video lectures outside of class, freeing up in class time for active learning (Tucker, 2012). Active learning has been found to produce better grades than passive learning (Minhas, Ghosh, & Swanzy, 2012). Collaborative learning takes place when two or more people learn something together, holding one another accountable for their learning (Roberts, 2004). Collaborative learning can create students who are more invested in their own learning, desiring to succeed in order to meet the expectations of one's peers (Roberts, 2004).

The Major Disadvantages of Inverted Instruction

From the review of literature, it was found that there are hundreds of articles and publications that refer to the flipped classroom, the teachers that use the method, or students' perceptions about it, but there is very little empirical data to quantify just how much students learn from the method (Hamdan, McKnight 2013). Within the limited data that exists, some contradicts each other. Arnold-Gaza (2013) and Nielson (2012) have negative perceptions towards the flipped classroom as they found that many students prefer the traditional classroom over the flipped classroom or do not have the appropriate tools at home to perform the flipped classroom. However, Gaughan (2013) concluded that the flipped classroom was successful in their experiment. Goodwin (2013) stated, "To date there is no scientific research base to indicate exactly how well flipped classrooms work."

In an extensive survey of research on the flipped classroom, which they explicitly defined as "an educational technique that consists of two parts: interactive group learning activities inside the classroom, and direct computer-based individual instruction outside the classroom," Bishop and Verleger (2013, p. 5) found only one empirical study, Day and Foley (2006), that examined student performance throughout a semester. Bishop and Verleger (2013) recommended that future research should objectively investigate student learning outcomes with controlled experimental designs and carefully consider the theoretical framework used in flipped classroom designs.

Although there are compelling reasons to implement an inverted classroom, there are also some potential problems. First, implementing an inverted classroom can initially be time-consuming. Teachers need to either carefully curate the videos from pre-made video sites or make their own videos (Flipped Learning Network, 2014). Both of these methods require an ample commitment of time from educators, and teachers must be prepared for the increased workload (Freeman Herreid & Schiller, 2013). An instructor cannot simply videotape a 50-minute lecture. Zappe, Leicht, Messner, Litzinger, and Lee (2009) found an optimum video length to be around 20 minutes, which requires the instructor to reorganize course material into short segments and to spend time editing recordings. The instructor must also develop and include activities to ensure that students are prepared for class (Day, Foley, 2006) (Kellog, 2009).

Second, online learning may frustrate some students. Strayer (2007) found some students were uncomfortable at having to take responsibility for their own learning. Students new to the method may be resistant initially because this new type of schooling requires them to do work at home rather than first be exposed to content and subject matter at school (Freeman Herreid & Schiller, 2013). The instructor can allay this discomfort by providing clear expectations for what students should know (Fredrickson, Reed, Clifford, 2005).

Third, there is some discrepancy in the literature about the appropriateness of an inverted classroom for different course levels. Some were cautious about using an inverted classroom in more advanced courses, while others suggest that an inverted classroom may be more applicable in advanced courses. (Baker,2000; Strayer, 2007).

The Major Misconceptions about Inverted Instruction

Some misconceptions about the flipped classroom are that student spend the entire time in front of a computer screen; students work without structure; videos replace the teacher; students work in isolation; or that a flipped classroom is an online course. I will address the major ones.

One misconception about the inverted classroom is that the flipped model is about replacing teachers with videos (Nochese, 2011). Some fear that the generation of online instructional videos will be used as a vehicle to weaken the role of teachers. One example critics point to is the Khan Academy, which is an archive of more than 4,000 videos made by Salman Khan, with the goal of changing education for the better by providing a free, world-class education to anyone anywhere (Khan, 2011). Critics have appropriately then questioned the need for teachers. Salman Khan has endorsed the flipped model and has stated that his videos allow the teacher to focus on higher-level learning activities, such as running simulations and labs with students, doing individual interventions, and facilitating peer-to-peer learning (Fink, 2011; Gojak, 2012). This emphasizes why the changes that occur in the classroom are the most important aspects of the inverted model. Bergmann and Sams argue that in a flipped learning environment, the role of teacher is amplified, in that all teachers now must know the individual learning needs of each student as their daily interactions increase. This actually increases the need for qualified, professional and caring educators. “Although video can be leveraged to deliver direct instruction, it does not, and cannot, replace the teacher as the facilitator of learning. (Bergman & Sams, 2012, p. 3)"

The second major misconception is that flipped learning is similar to an online course (Fink, 2011). Although online learning is – and will – -continue to have a valuable place in the education spectrum, it must be noted that an inverted model does not change the amount of face-to-face time that a student spends in a classroom when compared to a traditional classroom. However, the original definition of the flipped classroom – "what used to be classwork (the lecture) is done at home via teacher-created videos, and what used to be homework (assigned problems) is now done in class (Bergmann & Sams, 2012b)" – can imply that the flipped model may consist simply of online video lectures at home and a static use of class time for students to passively work on homework problems. This has led The Flipped Learning Network (2014) to release an updated and revised formal definition of flipped learning:

> Flipped Learning is a pedagogical approach in which direct instruction moves from the group learning space to the individual learning space, and the resulting group space in transformed into a dynamic, interactive learning environment where the educator guides students as they apply concepts and engage creatively in the subject matter. (p. 1)

Methods

The Background

STA 210 is a conceptual statistics course offered at University of Kentucky. It was created in 2010 following STA 200, a similar course that the department had taught for more than 25 years. Both were algebra-based courses that emphasized statistical concepts rather than mathematical manipulations. Euphemistically and somewhat incorrectly, the genre of courses like these has been labeled as "statistics for poets" around the country. Nonetheless, it is reasonable to house STA 210, just like STA 200, within liberal arts, requiring much more writing, reading, and conceptual ideas than a traditional statistics course.

Dr. Rayens created STA 210 for two reasons. First, the Department of Statistics was concerned that the conceptual statistics was too difficult (or at least too unfamiliar) for first-year teaching assistants (TAs) to present content and motivate students in the discussion breakouts, known as recitations. This concern prompted the department to move discussions and discoveries out of recitations and into classrooms. Second, there was too much passive learning in STA 200, based on recurring evidence on course final exams that many students were not retaining even a rudimentary understanding of important statistical concepts. This concern motivated the department to seek a new course in which less material was covered, and students needed to shoulder significantly more responsibility for learning the material.

The Department of Statistics at University of Kentucky has adopted a less-is-more approach for this course compared with what is typically seen as traditional content for this type of introductory conceptual statistics courses. Three modules of materials are included: human inference, confidence intervals, and hypothesis testing. Note that, as peculiar as this content focus may sound, the decision is purposeful for the course and more importantly does not distinguish the experimental conditions for this study.

There has been no textbook for the course, but students are required to purchase a workbook containing the content mentioned above. Dr. Rayens wrote the course workbook, *Beyond the Numbers: Student-Centered Activities for Learning Statistical Reasoning* (Van-Griner Publishing Company), currently in its fourth edition (in 2014). The workbook is structured around daily exercises (labeled as “Beyond the Numbers” or “BN”), larger projects (labeled as “Beyond the Classroom” or “BC”), and capstone projects (labeled as “CS”). BN activities are hands-on, designed to introduce important concepts, provide needed practices, and reflect on activities previously completed. Each of them covers a specific group of outcomes (see Appendices A and B). Materials for a host of activities (e.g., beer goggles for sensitivity and specificity testing) have been purchased and are available for use by STA 210 instructors and TAs from a keyed, common course supply room. Meanwhile, the content is recorded and placed on 18 videos. Since the fall of 2014, these videos have been available on YouTube for all students in the course, with specific instruction on how to access them in the workbook. Hence, all students in this experiment know about the videos and have free access to them. In other words, neither workbook nor videos differentiate the experimental conditions.

STA 210 became identified as “inverted” because lectures were removed from the classroom to make room for discussions and activities. With the traditional lecture time reimagined in this fashion, students were placed in an environment where they had to participate at some level when they came to class. This was further assisted by the creation of a workbook that consists of a series of prompts, applications, and hands-on activities that are designed to give expression and meaning to the reduced content. Of course, for students to benefit from the workbook, they need to actively engage in reading and completion.

The Experiment

At the University of Kentucky, approximately 4,000 undergraduates are taught statistical reasoning in the inverted classroom setting every year. The change from traditional to inverted classroom began in 2010 as an educational effort of Dr. William Rayens from the Department of Statistics to conduct a one-class pilot. Today, it has become a coordinated effort in more than 70 sections of the course STA 210 Introduction to Statistical Reasoning. This effort on average involves eight faculty instructors and approximately 20 TAs each calendar year. Study 2 will employ an educational experiment to investigate the impact of two different teaching methods for presenting statistical content on learning outcomes in introductory statistics courses. In this experiment, inverted instruction is the experiment condition (referred to as EXP) and traditional instruction is the control condition (referred to as business as usual or BAU). This controlled experiment aims to see if the inversion is making a measurable difference in cognitive and affective outcomes.

The inverted classroom environment at the University of Kentucky has been under construction for more than four years. Dr. Rayens created a series of ADA-compliant video lectures using Camtasia Studio and PowerPoint with audio narration, and these have been revised multiple times to their current professional status. The PowerPoints that backstop the videos were developed from transcripts that Dr. Rayens created. Transcripts and videos are available for reviewing (see Appendix C).

Participants in this experiment are approximately 135 students (blindly) enrolled in EXP and approximately 135 students (blindly) enrolled in BAU. That makes a total of 270 students as the sample for this experiment. Results of sample size and power analysis are summarized in Table 2. Effect size for standard calculations is the mean difference between EXP and BAU population means, scaled by population standard deviation assumed to be the same in both treatment populations. The table indicates that a sample size of 99 (per treatment group) is required to detect an effect size of 0.40 (in a two-tailed test) with 80 percent power for an alpha level of 0.05. This is a detection of one treatment outperforming the other by a four-tenths of a standard deviation. Cohen (1988) considered 0.20 as a small effect size, 0.50 as a medium effect size, and 0.80 as a large effect size. With 135 students in each of the EXP and BAU treatment groups, it is expected that this experiment can detect a near medium effect size with 80 percent power for the two-sample tests with two-tailed alternative or a small effect size with the same power for the two-sample tests with one-tailed alternative.

The experiment took place during the fall term of 2014 with undergraduate students who enrolled in STA 210. To reduce confronting effects, several measures were in place to have both EXP and BAU classes (a) offered to students simply as STA 210 so that students did not know prior to enrolling if they were taking the course in an inverted (EXP) or traditional (BAU) format (the University Registrar has confirmed that teaching styles do not need to be communicated to students prior to enrollment); (b) populated by students of similar gender, age, and ethnicity; (c) taught by the same instructor who worked with first-year TAs carefully trained to perform their respective roles in either EXP or BAU; (d) taught in similar physical classroom environments (i.e., typical lecture halls in the university's Whitehall Classroom Building); (e) taught at similar times of the day (daytime); (f) structured as large lectures enrolling approximately 135 students in a combined six sections with approximately 23 students in each section; (g) designed to meet two times each week with the full class and a third time in recitation for a total of three class hours each week (for each of the six sections); and (h) given the same assignments throughout the semester (could be in different settings and times). Condition (a) above serves as a kind of random sampling of students into the experimental groups to fulfill condition (b) above.

In sum, participants in this experiment were 135 students (blindly) enrolled in EXP and 135 students (blindly) enrolled in BAU group. That is a total of 270 students as the sample of this experiment. Demographic data were obtained from the university registrar's office to describe the students (see Table 1).

The Description of Common Components

Large lectures consist of six sections of approximately 24 students each. Every week, students meet together with the primary instructor for two hours and break out into a section-based recitation for the third hour. Medium (sized) lectures consist of three sections instead of six and are otherwise structured in the same way. This experiment worked with two large lectures based on six sections each. The entire course shares a common meeting structure. All sections spend two hours each week as one large class with the primary instructor (assisted by two first-year TAs) in a lecture hall that seats approximately 135 students, and each section spends one hour each week with a single TA in a recitation room that seats approximately 25 students. Time with the professor is often referred to as "main class time" and time with the TAs "recitation time."

In relation to the experiment, the time spent with the professor and TAs is exactly the same between EXP and BAU, and the physical spaces for those meetings to occur are exactly the same for the two experimental conditions. Specifically, in BAU, the main class time spent with the primary instructor is in a lecture format (i.e., PowerPoint presentations and videos) and the time spent with the TA for recitation is for discoveries and discussions. In EXP, discoveries and discussions from the recitation are brought into the classroom with the primary instructor for the two hours each week. The rote parts of the lecture (with PowerPoint presentations and videos) that used to occupy the primary instructor's time in the classroom are watched outside of class time. The recitations are used primarily to resolve some of the algebraic and proportional reasoning obstacles that often stand in the way of larger conceptual issues. Prior to those topics being discussed in a classroom activity, the instructor asks the TAs to work with students in recitation to make sure they can apply the needed computation and reasoning skills. Recitation activities like this are referred to as “fundamental practices,” in contrast to the original idea of using recitations for “discoveries and discussions.” A day-by-day explanation of activities in both experimental conditions is presented in detail for clarification in Appendices D and E.

The only variation between EXP and BAU pertains to the implementation of videos (video lectures). Videos are implemented in one large class (approximately 135 students) according to the way that the inverted method presents statistical content, and in the other large class (approximately 135 students) according to the way that the traditional method presents statistical content. This variation is precisely the consequences of the treatment (intervention) and thus can be considered a part of the treatment (intervention).

An extensive day-by-day comparison between EXP and BAU is constructed for the fall term of 2014 in Appendix E. Because how recitations take place depends on topics and treatments, the day-to-day work in BAU does not always match that in EXP. However, over a wider window of time, the same materials and assignments are implemented in both. In other words, in both EXP and BAU, the same topics are covered to the same depth, and students are expected to achieve the same level of mastery. Technology is not purposefully used as a part of treatment (intervention) but is applied when appropriate to enhance the teaching and learning within each environment (see Appendix F).

The Description of Experiment (EXP) Group

Students enrolled in the inverted environment (EXP) are required to watch videos (video lectures) before coming to class as part of their homework. When students come to class, they first see a short PowerPoint with the same content as the videos (about 10 minutes) and complete BN activities designed to help them engage with some of the complex concepts that were presented in the (content) videos. In some cases, the BN activities are homework as well, and the in-class activities provide parallel activities or discussions designed to bring out the subtleties of the BN activities for a better clarity of statistical concepts. In any case, students interact with each other and the instructor in main class time as they work to deepen their understanding of the material. Recitation directed by TAs is confined to fundamental practices (see discussion earlier and Appendices D and E). TAs in general do not facilitate active engagement with BN activities in recitation and avoid a selected group of BN activities (those marked with a star in Appendix D) because they are one of the main features of BAU. TAs answer questions, illustrate computations, distribute weekly quizzes created from the instructor, and supervise student completion of selected BN activities.

In sum, the main class time in EXP does not involve a traditional lecture. Summary comments and additional examples are provided in an opening PowerPoint presentation. The balance of the main class time is used for in-class exercises and discussions similar to those performed in recitations under the BAU treatment but so scaled that they are possible for the large class environment. Recitations are used to clarify numerical, procedural, and computational questions. Only a limited amount of active learning and conceptual discussion takes place in the recitation for the EXP treatment.

The Description of Business as Usual (BAU) Group

Students enrolled in the traditional environment (BAU) come to class two times a week and watch the same videos (video lectures) about the same statistical content. During lectures, the instructor uses the same PowerPoint (slides) to clarify concepts and definitions (note that, for both treatments, PowerPoints are the same in content and require the same amount of time in presentation). Traditional lectures are heavily content driven, where the instructor introduces statistical concepts and then works though examples to show the application of those concepts. As usual, these traditional lectures allow time for students to work though examples that apply concepts. During lectures, students have opportunities to ask questions and answer questions from the instructor related to examples under discussion. When appropriate, lectures are presented in an interactive manner. This environment is indeed a “business as usual” classroom that is lecture-focused or instructor-centered. Students also meet once a week in their individual recitation breakouts directed by first-year TAs. Recitations are where the primarily discovery-oriented BN activities are completed by students, with TAs who facilitate active engagement with BN activities. A selected group of BN activities must be facilitated and discussed by TAs in recitation (those marked with a star in Appendix D). TAs answer questions, illustrate computations, distribute weekly quizzes created by the instructor, and supervise student completion of selected BN activities.

In sum, the main class time in BAU is devoted to lectures and lecture-related activities (see Appendices D and E). A very limited amount of active learning and conceptual discussion take place in class. Instead, activities (leading to discoveries and discussions) are relegated to recitations with TAs.

The Variables

This study employs cognitive and affective measures as outcomes to examine treatment effects. Cognitive learning outcomes are defined as students' academic performance in statistics (with scores as indicators). These measures include (a) tests and exams, (b) completion of assignments (i.e., open-ended type of questions in homework or classwork during recitation time from the workbook) and (c) one or two major projects. Tests and exams contain both open-ended questions and multiple-choice questions. Students are given study guides with answers before tests and exams. Tests and exams measure performance in statistics in general and in identified conceptual constructs specifically. Conceptual constructs deemed critical to a basic statistical reasoning course have been developed under the direction and guidance of a faculty advisory committee from the Department of Statistics that considers it important to study how well students learn those constructs. A list of important items on which to measure competence is shown in Appendix G. Questions and activities are developed and identified to directly assess the entries on the list. Many come directly from the BN and BC activities that are already part of the course structure. Some of the questions are embedded in exams, and others are used in daily and recitation exercises.

Students' affective measures focus on attitude toward statistics (including interest, utility, motivation, and confidence) in the learning of statistics. Attitude, in general, is defined as "an individual's disposition to respond favorably or unfavorably to … any … discriminable aspect of the individual's world" (Ajzen, 1989, p. 241). Students' attitudes toward statistics refers to students' general impressions (i.e., positive or negative feelings) toward the discipline and learning of statistics as well as the way in which self is perceived in light of the practice (e.g., learning) of statistics (see Thurstone, 1970). Such a conception considers attitude toward statistics as a multidimensional construct of interest, utility, motivation, and confidence (anxiety) in the practice of statistics (see Organization for Economic Cooperation and Development, 2010). Specifically, interest refers to the level of enjoyment in the practice of statistics (e.g., liking or disliking statistics); utility refers to the usefulness, relevance, and value of statistics in life (i.e., personal and professional); motivation refers to the amount of effort that a student is inspired to spend on the practice of statistics; and confidence (anxiety) refers to the self-perception of competence in the handling of statistical knowledge and skills in an intellectual manner (Emmioglu & Capa-Aydin, 2012; Hood, Creed, & Neumann, 2012; Petocz & Newbery, 2010; Ramirez, Schau, & Emmioglu, 2012; Williams, 2013). Instruments measuring affective outcomes are administered as a survey (see Appendix H).

Independent variables portray student and course characteristics. Given that randomization is only partially achievable in this experiment, variables descriptive of student and course characteristics are important to control individual and practical differences between EXP and BAU and to examine whether they are able to enhance treatment effects. Student characteristics include gender, age, race, whether financial aid was provided (as a measure of SES), prior academic ability (i.e., ACT scores and overall GPA in high school), major at University of Kentucky, GPA of quantitative literacy courses taken at University of Kentucky, and cumulative GPA for the first and second years at University of Kentucky. Demographic and other useful information above can be collected from the university's student information system. Course characteristics include extra credit (e.g., bonus points), attendance for lectures and recitations, and usage of various learning materials (e.g., videos). As course characteristics, usage metrics from Courseload and click-through rates from Blackboard are collected to study the patterns of actual user interactions with videos and PowerPoints and to act as a statistical control to "purify" the treatment effects and to validate survey responses.

At the beginning of this semester-long experiment, students enrolled in the course are informed of the evaluation procedure and invited to take a survey about their experiences with statistics and the learning of statistics (affective measures). To ensure a good participation rate, the instructor provides incentives (extra credits) for participation and ensures that students sign consent forms. The survey is distributed after registration for the class is closed. During the last two weeks of the semester, students are invited to take the same survey plus some additions for EXP students to reflect on their experiences in the inverted environment (see Appendix H). Therefore, affective outcomes are measured in a pretest and posttest fashion. Cognitive measures (statistics performance) are obtained mostly during the second half of the semester and are obviously considered posttest measures.

Student surveys are distributed using an online survey software, Qualtrics. An adaptive release mechanism is used to hide the survey until after students digitally sign the consent form by reading the form and answering a question to confirm their understanding of the procedure. Students' university IDs are collected to link the two surveys together for comparison.

The Analysis

Multiple regression/correlation (MRC) analysis are used to test the between-group differences in cognitive and affective learning outcomes between students in inverted and traditional classrooms. Specifically, MRC is used to compare statistics performance in general and in the identified conceptual constructs in particular for the EXP and BAU conditions with control of student and course characteristics, particularly prior ability in quantitative literacy (e.g., ACT math score or GPA of quantitative literacy courses taken at University of Kentucky). Because affective measures were obtained in a pretest and posttest manner, analysis of covariance (ANCOVA) in the form of MRC is used to compare affective measures between the EXP and BAU conditions also with control of student and course characteristics. MRC also is used to examine students in the inverted classroom by linking within-group differences in cognitive and affective learning outcomes to variables descriptive of student and course characteristics in order to identify salient student and course characteristics that enhance cognitive and affective learning outcomes in EXP.

Table 3.1.1

Descriptive Statistics (Percentages) of Demographics from Fall 2013

Variable	Category	Morning	Afternoon	Evening
Age	Teens	53.79	47.50	45.15
	Early 20s	42.84	49.44	47.57
	Late 20s	2.29	1.67	4.13
	Over 30	1.08	1.39	3.16
Gender	Male	43.92	48.06	58.50
	Female	56.08	51.94	41.50
Race	Black	9.63	7.50	6.80
	White	77.74	79.72	75.49
	Asian	3.97	3.61	5.83
	Other	8.66	9.17	11.89

Table 3.1.2

Summary of Power Study

		Two-Sample t Test	
Effect Size	Paired t Test	One-Tailed	Two-Tailed
0.1	620	1238	1570
0.2	156	310	393
0.3	71	139	175
0.4	41	78	99
0.5	27	51	63
0.6	19	36	44
0.7	15	26	33
0.8	12	21	25
0.9	10	16	20
1.0	8	14	16
2.0	4	4	5

RESULTS

Descriptive Statistics

As stated earlier, this study investigated the impact of different teaching methods for presenting statistical content (i.e., inverted classroom versus traditional classroom) on learning outcomes in introductory statistics. According to Table 1, participants are 135 students (blindly enrolled) in EXP and 135 students (blindly enrolled) in BAU. That is a total of 270 students as the sample for this experiment.

Table 1 presents descriptive statistics of the seven outcome variables between the two treatment conditions, overall projects average, overall tests average, overall classwork, midterm attendance average, class final attendance average, midterm grade and class final grade. The table shows 3.61 points difference in the overall projects average score in favor of the BAU group, 1.35 points difference in overall tests average in favor of the BAU group, 4.95 points difference in overall classwork in favor of the BAU group, 3.65 points difference in midterm attendance average in favor of the BAU group, 3.47 points difference in class final attendance average in favor of the BAU group, 3.21 points difference in midterm grades in favor of the BAU group, and 3.03 points difference in class final grade in favor of the BAU group.

As predictor variables, we used available individual student characteristics that we grouped into three different blocks. The first block includes individual background variables of gender, age and ethnicity. There are 59% males in the BAU group and 46% males in the EXP group. Age for both groups was very similar with M=20.41, SD=0.81 for the BAU group versus M=20.72, SD=1.38 for the EXP group. Ethnicity is also very similar in both groups with 78% white students in the BAU group and 72% white students in the EXP group.

The second block includes high school background variables of high school grade point average (GPA) and ACT mathematics score. The high school ACT mathematics score for both groups is also very similar with M=24.68, SD=3.88 for the BAU group versus M=24.39, SD=4.38 for the EXP group. The high school GPA for both groups is also very similar with M=3.53, SD=0.68 for the BAU group versus M=3.47, SD=0.76 for the EXP group.

The last block included university program-based variables of major and cumulative GPA. The university cumulative GPA shows a very small advantage for the BAU group with M=3.11, SD=0.61 versus M=3.02, SD=0.59 for the EXP group. As for the six university majors, 5% of the students in the BAU group majored in engineering, compared with 15% of the students in the EXP group, 24% of the students in the BAU group majored in education and nursing professional fields, compared with 16% of the students in the EXP group, 20% of the students in the BAU group majored in economics, compared with 22% of the students in the EXP group, 18% of the students in the BAU group majored in humanities, compared with 21% of the students in the EXP group, 15% of the students in the BAU group majored in sciences, compared with 7% of the students in the EXP group, 17% of the students in the BAU group declared an undecided major, compared with 20% of the students in the EXP group.

Comparison of Treatment with Student Characteristics

In Table 2, seven outcome variables were examined by means of a hierarchical regression analysis from the perspective of student background as various blocks (e.g., Block 1 pertains to student age, gender and ethnicity). There were seven different outcome (dependent) variables: overall projects average (based on three projects), overall tests average (based on three tests), overall classwork (based on homework, in-class assignments and quizzes), midterm attendance average, class final attendance average, midterm grade and class final grade.

The block of student background variables included age, gender, and ethnicity (functioned as control independent variables). Student age was centered around its grand mean. The block of treatment included the key independent variable, the treatment dummy, comparing inverted EXP with traditional BAU. Overall, the model accounted for 9% of the variance in projects average (statistically significant), 1% in tests average, 3% in overall classwork (statistically significant), 4% in midterm attendance average, 3% in class final attendance average, 10% in midterm grade (statistically significant) and 4% in class final grade.

The R^2 change indicated the relative importance of each block to a certain outcome measure. The block of student background variables (age, gender, and ethnicity) had a R^2 change of 0.08 in projects average (statistically significant) compared with the (block of) treatment that had a R^2 change of 0.02 (statistically significant), 0.02 in tests average compared with 0.00, 0.03 in overall classwork compared with 0.02 (statistically significant), 0.05 in midterm attendance average (statistically significant) compared with 0.01, 0.04 in class final attendance average compared with 0.01, 0.10 (statistically significant)in midterm grade compared with 0.02 (statistically significant) and finally, class final grade 0.04(statistically significant) compared with 0.01. Across the outcome measures, the block of student background variables is much more important than the (block of) treatment.

The block of student high school background variables included high school grade point average (GPA) and ACT mathematics score (functioned as control independent variables). Both variables— high school GPA and ACT mathematics score— were centered around its grand mean. The block of treatment included the key independent variable, the treatment dummy, comparing inverted EXP with traditional BAU. Overall, the model accounted for 9% of the variance in projects average (statistically significant), 44% in tests average, 4% in overall classwork (statistically significant), 3% in midterm attendance average, 2% in class final attendance average, 21% in midterm grade (statistically significant) and 24% in class final grade.

The R^2 change indicated the relative importance of each block to a certain outcome measure. The block of student high school background variables (high school GPA and ACT mathematics score) had a R^2 change of 0.09 in projects average (statistically significant) compared with the (block of) treatment that had a R^2 change of 0.02 (statistically significant), 0.45 (statistically significant) in tests average compared with 0.00, 0.03 in overall classwork (statistically significant) compared with 0.02 (statistically significant), 0.04 in midterm attendance average (statistically significant) compared with 0.01, 0.02 in class final attendance average compared with 0.01, 0.20 (statistically significant) in midterm grade compared with 0.02 (statistically significant) and finally, class final grade 0.25 (statistically significant) compared with 0.01. Across the outcome measures, the block of student high school background variables is much more important than the (block of) treatment.

The block of student university program background variables included student major and university cumulative GPA (functioned as control independent variables). University cumulative GPA was centered around its grand mean. The block of treatment included the key independent variable, the treatment dummy, comparing inverted EXP with traditional BAU. Overall, the model accounted for 33% of the variance in projects average, 52% in tests average, 27% in overall classwork, 16% in midterm attendance average, 20% in class final attendance average, 49% in midterm grade and 57% in class final grade.

The R^2 change indicated the relative importance of each block to a certain outcome measure. The block of student university program background variables (student major and university cumulative GPA) had a R^2 change of 0.34 in projects average (statistically significant) compared with the (block of) treatment that had a R^2 change of 0.00, 0.53 in tests average (statistically significant) compared with 0.00, 0.29 in overall classwork (statistically significant) compared with 0.01, 0.19 in midterm attendance average (statistically significant) compared with 0.00, 0.22 in class final attendance average (statistically significant) compared with 0.00, 0.50 (statistically significant)in midterm grade compared with 0.01 and finally, class final grade 0.58 (statistically significant) compared with 0.00. Across the outcome measures, the block of student background variables is much more important than the (block of) treatment.

The last model included all three student background blocks together (functioned as control independent variables): individual background, high school background and university program background. The block of student background variables included age, gender, and ethnicity. The block of student high school background variables included high school GPA and ACT mathematics score. The block of student university program background variables included student major and university cumulative GPA. The block of treatment included the key independent variable, the treatment dummy, comparing inverted EXP with traditional BAU. Overall, the model accounted for 38% of the variance in projects average, 60% in tests average, 35% in overall classwork (statistically significant), 19% in midterm attendance average, 23% in class final attendance average, 54% in midterm grade (statistically significant) and 65% in class final grade.

The R^2 change indicated the relative importance of each block to a certain outcome measure. In terms of projects average, the block of student background variables had a R^2 change of 0.07 (statistically significant), the block of student high school background variables had a R^2 change 0.06 (statistically significant), the block of student university program background variables had a R^2 change of 0.28 (statistically significant), and the (block of) treatment that had a R^2 change of

0.01. In terms of test average, the block of student background variables had a R^2 change of 0.02, the block of student high school background variables had a R^2 change 0.45 (statistically significant), the block of student university program background variables had a R^2 change of 0.16 (statistically significant), and the (block of) treatment that had a R^2 change of 0.00. In terms of classwork, the block of student background variables had a R^2 change of 0.02, the block of student high school background variables had a R^2 change 0.02, the block of student university program background variables had a R^2 change of 0.33 (statistically significant), and the (block of) treatment that had a R^2 change of 0.02 (statistically significant). In terms of midterm attendance average, the block of student background variables had a R^2 change of 0.04 (statistically significant), the block of student high school background variables had a R^2 change 0.01, the block of student university program background variables had a R^2 change of 0.17 (statistically significant), and the (block of) treatment that had a R^2 change of 0.01. In terms of class final attendance average, the block of student background variables had a R^2 change of 0.03, the block of student high school background variables had a R^2 change 0.01, the block of student university program background variables had a R^2 change of 0.23 (statistically significant), and the (block of) treatment that had a R^2 change of 0.01. In terms of midterm grade, the block of student background variables had a R^2 change of 0.08 (statistically significant), the block of student high school background variables had a R^2 change 0.18 (statistically significant), the block of student university program background variables had a R^2 change of 0.30 (statistically significant), and the (block of) treatment that had a R^2 change of 0.01 (statistically significant). In terms of class final grade, the block of student background variables had a R^2 change of 0.03, the block of student high school background variables had a R^2 change 0.24 (statistically significant), the block of student university program background variables had a R^2 change of 0.40 (statistically significant), and the (block of) treatment that had a R^2 change of 0.01 (statistically significant). Across the outcome measures, each block of student characteristics turned out to be much more important than the (block of) treatment.

Table 3 presents the results for the absolute treatment effects of inverted classroom against traditional classroom in terms of projects average, tests average, classwork, midterm attendance average, class final attendance average, midterm grade and class final grade. The results showed a statistically significant treatment effect on projects average (Effects=-3.61, SE=1.50), classwork (Effects=-4.95, SE=1.98), midterm attendance average (Effects=-3.65, SE=1.80), midterm grade (Effects=-3.21, SE=1.05) and class final grade (Effects=-3.04, SE=1.42), all in favor of students in traditional classroom group (BAU). Therefore, students in the traditional classroom did better than students in the inverted classroom in projects average, classwork, midterm attendance average, midterm grade and class final grade.

Table 4 presents the simplified results for the relative treatment effects of inverted classroom against traditional classroom in terms of projects average, tests average, classwork, midterm attendance average, class final attendance average, midterm grade and class final grade, with the control of student background variables (gender, age and ethnicity). The results showed a statistically significant treatment effect on projects average (Effects=-3.02, SE=1.35), classwork (Effects=-3.95, SE=1.85), and midterm grade (Effects=-2.19, SE=1.05), all in favor of students in traditional classroom group (BAU). Therefore, after controlling for student background variables of gender, age and ethnicity, students in the traditional classroom did better than students in the inverted classroom in projects average, overall classwork and midterm grade.

Table 5 presents the simplified results for the relative treatment effects of inverted classroom against traditional classroom in terms of projects average, tests average, classwork, midterm attendance average, class final attendance average, midterm grade and class final grade, with the control of student high school background variables (high school grade point average and ACT mathematics score). The results showed a statistically significant treatment effect on projects average (Effects=-2.77, SE=1.33), classwork (Effects=-4.08, SE=2.00), and midterm grade (Effects=-2.26, SE=1.04), all in favor of students in traditional classroom group (BAU). Therefore, after controlling for student high school background variables of high school GPA and ACT mathematics score, students in the traditional classroom did better than students in the inverted classroom in projects average, overall classwork and midterm grade.

Table 6 presents the simplified results for the relative treatment effects of inverted classroom against traditional classroom in terms of projects average, tests average, classwork, midterm attendance average, class final attendance average, midterm grade and class final grade, with the control of student university background variables (major and university cumulative GPA). The results showed a statistically not significant treatment effect on all outcomes measured. Therefore, after controlling for student university background variables of student major and university cumulative GPA, students in the traditional classroom did similar to students in the inverted classroom in projects average, test average, overall classwork, midterm attendance average, class final attendance, midterm grade and class final grade.

Table 7 presents the simplified results for the relative treatment effects of inverted classroom against traditional classroom in terms of projects average, tests average, classwork, midterm attendance average, class final attendance average, midterm grade and class final grade, with the control of student background variables (gender, age and ethnicity), high school background variables (high school GPA and ACT mathematics score) and university program background (student major and university cumulative GPA). The results showed a statistically significant treatment effect on midterm grade (Effects=-1.69, SE=0.82), in favor of students in traditional classroom group (BAU). Therefore, after controlling for student background variables, student high school background variables and university program variables students in the traditional classroom did better than students in the inverted classroom in midterm grade only.

Table 3.1

Descriptive Statistics of Outcome and Predictor Variables

	BAU (n = 130)		EXP (n =135)	
	Mean	SD	Mean	SD
Outcome Variables (Dependent Variables)				
Projects Average (continuous)	87.04	10.87	83.43	13.40
Tests Average (continuous)	80.11	12.64	78.76	11.57
Classwork (continuous)	87.28	16.34	82.33	15.93
Midterm Attendance Average (continuous)	89.85	13.61	86.20	15.56
Class Final Attendance Average (continuous)	88.88	14.53	85.41	15.71
Midterm Grade (continuous)	86.47	8.36	83.26	8.70
Class Final Grade (continuous)	85.96	11.98	82.93	11.07
Predictor Variables (Independent Variables)				
Male (= 1, female = 0)	0.59	0.49	0.46	0.50
Age (continuous)	20.41	0.81	20.72	1.38
Ethnicity (white = 1, nonwhite = 0)	0.78	0.42	0.72	0.45
High School ACT mathematics score (continuous)	24.68	3.88	24.39	4.38
High School GPA (continuous)	3.53	0.68	3.47	0.76

University Cumulative GPA (continuous)	3.11	0.61	3.02	0.59
Major 1 Engineering (yes = 1, no = 0)	0.05	0.23	0.15	0.36
Major 2 Professional (Education, Nursing) (yes = 1, no = 0)	0.24	0.43	0.16	0.36
Major 3 Economics (yes = 1, no = 0)	0.20	0.40	0.22	0.42
Major 4 Humanities (yes = 1, no = 0)	0.18	0.39	0.21	0.41
Major 5 Sciences (yes = 1, no = 0)	0.15	0.36	0.07	0.25
Major 6 Undecided (yes = 1, no = 0)	0.17	0.38	0.20	0.40

Table 3.2

R Square Change and Proportion of Variance Explained in Various Hierarchical Regression Models Examining Treatment Effects (Inverted Classroom versus Traditional Classroom)

	Proportion of Variance Explained	R^2 Change for Block 1	R^2 Change for Block 2	R^2 Change for Block 3	R^2 Change for Block 4
Block 1 = Individual Background, Block 2 = Treatment Condition					
Projects Average	0.09*	0.08*	0.02*		
Tests Average	0.01	0.02	0.00		
Classwork	0.03*	0.03	0.02*		
Midterm Attendance Average	0.04	0.05*	0.01		
Class Final Attendance Average	0.03	0.04	0.01		
Midterm Grade	0.10*	0.10*	0.02*		
Class Final Grade	0.04	0.04*	0.01		
Block 1 = High School Background, Block 2 = Treatment Condition					
Projects Average	0.09*	0.09*	0.02*		
Tests Average	0.44	0.45*	0.00		
Classwork	0.04*	0.03*	0.02*		
Midterm Attendance Average	0.03	0.04*	0.01		
Class Final Attendance Average	0.02	0.02	0.01		
Midterm Grade	0.21*	0.20*	0.02*		
Class Final Grade	0.24	0.25*	0.01		
Block 1 = University Program Background, Block 2 = Treatment Condition					
Projects Average	0.33	0.34*	0.00		
Tests Average	0.52	0.53*	0.00		
Classwork	0.27	0.29*	0.01		

Midterm Attendance Average	0.16	0.19*	0.00		
Class Final Attendance Average	0.20	0.22*	0.00		
Midterm Grade	0.49	0.50*	0.01		
Class Final Grade	0.57	0.58*	0.00		
Block 1 = Individual Background, Block 2 = High School Background, Block 3 = University Program Background, Block 4 = Treatment Condition					
Projects Average	0.38	0.07*	0.06*	0.28*	0.01
Tests Average	0.60	0.02	0.45*	0.16*	0.00
Classwork	0.35*	0.02	0.02	0.33*	0.02*
Midterm Attendance Average	0.19	0.04*	0.01	0.17*	0.01
Class Final Attendance Average	0.23	0.03	0.01	0.23*	0.01
Midterm Grade	0.54*	0.08*	0.18*	0.30*	0.01*
Class Final Grade	0.65*	0.03	0.24*	0.40*	0.01*

Note. The block of individual background includes gender, age, and race-ethnicity. The block of high school background includes high school GPA and ACT mathematics score. The block of university program background includes student major and university cumulative GPA.

Table 3.3

Results of Hierarchical Regression Analysis on Absolute Treatment Effects (Inverted Classroom versus Traditional Classroom)

	Effects	SE
Projects Average	-3.61*	1.50
Tests Average	-1.36	1.49
Classwork	-4.95*	1.98
Midterm Attendance Average	-3.65*	1.80
Class Final Attendance Average	-3.47	1.86
Midterm Grade	-3.21*	1.05
Class Final Grade	-3.04*	1.42

* $p < .05$.

Table 3.4

Results of Hierarchical Regression Analysis on Relative Treatment Effects (Inverted Classroom versus Traditional Classroom) with Control of Individual Background

	Effects	SE
Projects Average	-3.02*	1.35
Tests Average	-.35	1.56
Classwork	-3.95*	1.85
Midterm Attendance Average	-2.88	1.88
Class Final Attendance Average	-2.63	1.84
Midterm Grade	-2.19*	1.05
Class Final Grade	-2.15	1.30

Note. Treatment effects are adjusted over the block of individual background including gender, age, and race-ethnicity.

* $p < .05$.

Table 3.5

Results of Hierarchical Regression Analysis on Relative Treatment Effects (Inverted Classroom versus Traditional Classroom) with Control of High School Background

	Effects	SE
Projects Average	-2.77*	1.33
Tests Average	0.56	1.13
Classwork	-4.08*	2.00
Midterm Attendance Average	-2.71	2.12
Class Final Attendance Average	-2.68	2.04
Midterm Grade	-2.26*	1.04
Class Final Grade	-1.62	1.15

Note. Treatment effects are adjusted over the block of high school background including high school GPA and ACT mathematics score.

* $p < .05$.

Table 3.6

Results of Hierarchical Regression Analysis on Relative Treatment Effects (Inverted Classroom versus Traditional Classroom) with Control of University Program Background

	Effects	SE
Projects Average	-1.43	1.22
Tests Average	1.10	1.09
Classwork	-2.08	1.74
Midterm Attendance Average	-1.84	1.81
Class Final Attendance Average	-1.47	1.74
Midterm Grade	-1.31	0.82
Class Final Grade	-0.51	0.92

Note. Treatment effects are adjusted over the block of university program background including student major and university cumulative GPA.

* $p < .05$.

Table 3.7

Results of Hierarchical Regression Analysis on Relative Treatment Effects (Inverted Classroom versus Traditional Classroom) with Control of Individual Background, High School Background, and University Program Background

	Effects	SE
Projects Average	-1.81	1.17
Tests Average	0.75	1.00
Classwork	-3.97	1.65
Midterm Attendance Average	-2.65	2.10
Class Final Attendance Average	-2.68	1.87
Midterm Grade	-1.69*	0.82
Class Final Grade	-1.37	0.80

Note. Treatment effects are adjusted over the blocks of individual background (gender, age, ethnicity), high school background (high school GPA and ACT mathematics score), and university program background (student major and university cumulative GPA) in this order.

* $p < .05$.

DISCUSSION

Summary of Principal Findings

This research studied use of the flipped classroom in undergraduate statistics class and its effect on student achievement. Students in the BAU group received traditional lecture, and students in the experimental group received inverted lecture. We compared seven outcomes for the two groups: projects average, tests average, classwork, midterm attendance average, class final attendance average, midterm grade and class final grade. We used three different blocks with student background variables as predictors. For the first one, individual student background included age, gender and ethnicity. For the second one, high school background variables included high school GPA and ACT mathematics scores. In the third one, university program background included university cumulative GPA and student majors. Without any control over student characteristics (i.e., for the absolute treatment effects of inverted classroom against traditional classroom), the results show that students in the traditional classroom did better than students in the inverted classroom in projects average, classwork, midterm attendance average, midterm grade and class final grade.

After controlling for student background variables of gender, age and ethnicity, students in the traditional classroom did better than students in the inverted classroom in projects average, overall classwork and midterm grade. The model, when controlling for student high school background variables of high school GPA and ACT mathematics score, showed that students in the traditional classroom did better than students in the inverted classroom in projects average, overall classwork and midterm grade. Finally, after controlling for student university background variables of student major and university cumulative GPA, students in the traditional classroom performed similarly to students in the inverted classroom in projects average, test average, overall classwork, midterm attendance average, class final attendance, midterm grade and class final grade. When controlling for all, student background variables—student high school background variables and university program variables—students in the traditional classroom did better than students in the inverted classroom in midterm grade only.

Insights to Research Literature

Education literature suggests that inverted classrooms support active learning, and this may benefit students more than a traditional lecture-type model (Berrett, 2012; Bergman & Sams, 2012). This is because students in an inverted class perform the lower-order, easier tasks from Bloom's revised taxonomy (Anderson et al., 2005) outside of class and the higher-order, more difficult tasks in class, with instructor and peer support. Active learning is more difficult than passive learning, but the payoff is potentially greater because activities cement concepts in students' minds more permanently than if students only read the material. (Touchton, 2015) This study thus compared the effectiveness of using traditional lecture format and inverted classroom format in undergraduate level statistics class. We found that inverted and traditional classroom are equally effective academically across all our outcome measures except one (i.e., midterm grade). This exception may be easy to reason. Given enough time (i.e., a whole semester instead of a half semester), there would be no differences between the two groups of students.

The findings of the present study do not offer direct support to the claim that inverting the classroom creates an environment emphasizing goal-directed practice and feedback that would improve learning outcomes (e.g., Ambrose et al., 2010; Mason et al., 2015). Nonetheless, considering the fact that an inverted classroom is much more difficult to create, operate, and maintain, the non-significant findings of the present study may actually be good news, suggesting at the very least that the inverted classroom, given all its difficulties, can be just as effective as the traditional classroom. This may be the first step, the cornerstone, toward building statistically significant advantages of the inverted classroom in the near future when improvements can be made in terms of curriculum and instruction as well as operation and management for the inverted classroom.

Along this line of thinking, the present study seems to suggest that students would need time to adjust to the inverted classroom format. Students in the experimental group may be a bit off balance concerning the way their class is run. Frederickson, Reed, and Clifford (2005) showed that students in a technology-rich environment, where the professor is less visible, require different things than students in a traditional lecture course. One of their significant results stated that students learning with technology need more reassurance that they were “on the right track and doing the right thing” during the learning process. This suggests that class rules, division of labor, and structure of the community are all significantly affected (and changed) when students use a different major tool (technology and inverted format class) to learn content, when compared to a traditional lecture style class. This perspective could suggest ways of improvement towards a better and more effective inverted classroom.

Setting this study in an undergraduate level introductory statistics large class makes it unique from other studies in the literature in a few important respects. An argument can be made that the inverted classroom is a more natural fit for some topics and a less natural fit for others. The Frederickson, Reed, and Clifford (2005) study was set in a statistics course, but it was at the graduate level. The success of the study suggests that the flip format may work best in a setting where most of the students in the course are deeply interested in the content. Students in this position would be motivated to take it upon themselves to do what it takes outside of class so they will be productive during activities inside the classroom, the authors explained. In Frederickson et al. (2005), students in the inverted classroom and the traditional groups both performed at the same level, but students in the flipped classroom had concerns about the structure of the classroom. That graduate level students struggled with adjusting to the flipped classroom format further suggests that an introductory course at the undergraduate level may face more challenges in implementing the inverted classroom. In such an introductory statistics course, it is possible that students come in wanting to be introduced to the subject rather than expected to devise their own ways of thinking about the subject. This is a practical dynamic of introductory courses that cannot be ignored.

Implications for Educational Policies and Practices

Because traditional and inverted classroom are equally effective, educational authorities may not need to promote inverted instruction as a major educational reform. Instead, efforts to promote research and development to engage in the improvement of the inverted classroom may become priority. Some efforts seem obvious. For example, in the present study, the teaching assistants who helped with the inverted classroom group did not have any training on the tenets of the flipped classroom. Therefore, the inverted classroom group did not receive some of the benefits of the flipped classroom model designed to create a dynamic, inquiry-based learning environment. The students did not have the full advantage of working problems in collaborative groups.

For another example, the instructor is highly expected to use recitation time for the inverted classroom group as an active session where students work collaboratively in groups, present statistics problems to the class or in their groups and have active whole-class discussions (not lectures). In the present study, a lack of instructor training of these instructional methods might be responsible for insufficient implementation of these inquiry-based and collaborative learning techniques. The instructor needs to be a content expert and a pedagogical expert (and sometimes even a classroom manager), all at the same time.

One of the reasons we see no treatment effects is the fact that Inverted Classroom (IC) was a new instructional method for students and they may need some time to get used to it. Most students are instructed in a teacher-center fashion through all of their education. These students may have confusion and even reluctance to engage in the inverted instruction. Although this is speculation, there seems to be a need for the teaching staff to recognize and work with this potential hurdle when students walk into an inverted classroom.

Some students may also be skeptical at the beginning of the method with such comments as that the professor is expecting them to teach themselves. Learning the content by themselves as homework and getting used to the hands-on learning activities during class time may seem too much responsibility for students, suggesting a diminished role of the instructor. This, again, speaks to the need for the teaching staff to help students understand and appreciate the real intention of this instructional format.

Both potential reasons may be valid given the fact that the only significant effects in favor of students in traditional format occurred at midterm (i.e., midterm grade), which disappeared at the end of the course. It seems that when student really get into the method, things start to take a positive turn. When students experience the instructional approach as motivating, engaging and unique, they may begin to engage. To help make this transition happen, a blended approach of lecturing and active learning in class through flipping may be appropriate. For example, it may help if the teaching staff practices the strategy to free up class time using videos (rather than flipping every class).

Because the videos that were created for the flipped class were made available also to the students in the control sections, students in the BAU sections might have some advantage in terms of learning resources. Although it is not clear at this time what differences this addition might have made, it might explain the lack of statistically significant treatment effects.

Limitation and Suggestion for Further Research

Nowadays with increased use of Internet technology, virtual communications and learning management systems, many university instructors are interested in inverted or flipped classrooms. The purpose of this study was to compare the effectiveness of a traditional lecture-type classroom with the inverted classroom when teaching statistics to undergraduate students in a core content course. The results of the study may not be generalized to the population of all undergraduate students. It also gives no indication as to how the results would generalize to other content domains. Further studies may explore along these lines of inquiry regarding the effects of the traditional classroom in comparison with inverted.

The lack of statistically significant advantages of the inverted classroom over the traditional classroom may, to a large extent, indicate that with only one semester to compare the two teaching methods in this study, the duration may not be sufficient. Further studies may seek some longer period of using and comparing the two teaching methods.

Although not perfect measures, outcome measures used in the present study were considered reasonable, valid and diverse assessments of subject mastery. There is one limitation in using these outcome measures in that scoring or grading on most outcome measures is subjective. Standardized assessment tools have a critical role to play in further research on the inverted classroom.

One way to ensure the discovery of the advantages of the inverted classroom may be the standardization of the experimental procedures. In the present study, different teaching assistants were used for the two different groups, even though all assistants were first-year graduate students with similar individual backgrounds. There was really no control over the behaviors and efforts of these teaching assistants. This situation contributes to the data and is left to future researchers to resolve with increased standardization of the experimental procedures.

Finally, it is worth emphasizing what perhaps is the largest limitation in the present study—the lack of pretest data. This lack could be a major reason why the advantages of the inverted classroom were unseen in the present study. A pretest-posttest research design is highly desirable to investigate the concept of the inverted classroom. For various administrative and practical difficulties, the present study did not adopt a pretest-posttest research design, but it is strongly recommended that further research along this line of inquiry adopts such a research design. Very often, the final status in a semester is not really informative or important, but the growth or change during the semester is what matters most.

CHAPTER 4: CONCLUSION

Motivation for Educational Experiments

The ultimate goal of statistics education is to create a statistically literate society in which people can appropriately use statistical thinking (Kettenring, Lindsay, & Siegmund, 2004; Schau, 2003). This dissertation research comes at a historical time when there is a strong emphasis on the need to improve students' ability to think statistically at all educational levels. There is a growing movement to introduce concepts of statistics and probability into the elementary and secondary mathematics curriculum, and there are calls for teaching statistics and probability in a deeper and different way than has been done (NCTM, 2000; CCSSM, 2010). Educational reforms in K-12 mathematics education are creating considerable impact on undergraduate statistics education at the college level. Although the need to improve the teaching of introductory statistics courses is not a new one, with increased demand on these courses, there has been constant effort to seek out better ways of teaching these courses (e.g., Garfield, Hogg, Schau, & Whittinghill, 2002; Lindsay et al., 2004).

In recent years, many statisticians have become involved in the ongoing reform of the teaching of introductory statistics (e.g., Garfield & Ben-Zvi, 2008), and the National Science Foundation has funded numerous projects in promotion of this reform (e.g., Garfield et al., 2002). The University of Kentucky (UK) began a reform of its general education program in November 2005 and formally implemented the new General Education Program in May 2009 (often referred to as UKCore) (see http://www.uky.edu/ukcore/). Thinking and reasoning are the central themes of this well-designed general education curriculum. In sum, the stakes for reform in mathematics and statistics education are high not only in public schools but also in colleges and universities (Klein, 2003).

Many research studies over the past several decades indicate that most students and adults cannot think statistically about important issues that affect their lives (Garfield & Ben-Zvi, in press), even though their lives are increasingly governed by numbers (Moore, 1997). Tishkovskaya and Lancaster (2012) argued that our society has entered into an age of information where the "information explosion" is creating a critical need for statistically educated citizens — people who need to be statistically literate not only in their workplace but also in their everyday lives.

Addressing the need to improve students' ability to think statistically, schools are making statistical reasoning a critical part of the mainstream mathematics curriculum around the world (Batanero, Burill, & Reading, 2011). Statistics education is critical in today's data-rich economy because it can promote the "must-have" competencies essential to "thrive in the modern world" (Franklin et al., 2007, p. 4).

Research literature is full of students' inabilities to understand statistical concepts and procedures, a strong indication of the need for reform in statistics education. Ben-Zvi and Garfield (2004) discussed some of the reasons that explain why statistics is a challenging discipline to learn and to teach. First, many statistical ideas and rules are complex, difficult, and even counterintuitive so as to discourage students to engage in the learning of statistics. Second, many students have difficulty with the underlying mathematics (e.g., fractions, decimals, proportional relationship, algebraic manipulation), which interferes with the learning of statistical concepts and procedures. Third, the context in many statistical problems tends to mislead students to rely on experiences and often faulty intuitions to produce a solution rather than select an appropriate statistical procedure and rely on data-generated evidence. A fourth reason is that students equate statistics with mathematics and expect the focus to be on numbers, computations, and formulas, all leading to just one correct answer. Finally, inadequate experiences fail to prepare students for the massiveness of data, the different possible interpretations based on different assumptions, and the extensive reliance on communication skills.

With the separation of statistics from mathematics, statistics educators are still trying to fully understand the challenges and difficulties in teaching and learning statistics as a unique discipline. Reforms in statistics education is ongoing. Statistics educators over the last decade have called for the development of statistical literacy and interpretive skills as the universal goals of statistics education (e.g., Del Mas, 2002; Rumsey, 2002). Part of this reform seeks for better alignment of instruction with important learning goals and assessments (Garfield & Gal, 1999).

Despite a growing body of research related to the teaching and learning of statistics at all educational levels, few direct connections have been established between research and practice (Garfield & Zvi, 2009). The main goal of this dissertation is to fill in some gaps in the research literature on the teaching and learning of statistics. By nature, this dissertation joins the reform effort of shift in content and pedagogy as discussed earlier. To promote the link between research and practice, educational experiments are used to examine the effects on learning outcomes of different instructional practices in statistics education, in particular the use of different types of manipulatives and different styles of instruction. Specifically, this dissertation includes two independent studies (experiments). The first study examines the instructional effects of physical versus virtual manipulatives (see definitions later) on learning outcomes in introductory statistics, whereas the second study investigates the impact of different styles in teaching statistics (inverted classroom versus traditional classroom) on learning outcomes in introductory statistics. The results of these studies will improve undergraduate statistical education and provide meaningful links between research and practice. In general, this dissertation strives to join many other reform efforts to explore instructional ways that engage students in reasoning and thinking statistically.

To combat the abstract nature of probability and statistics, the use of manipulatives may represent one of the most effective strategies in the statistics classroom. Manipulatives enhance the abilities of students at all levels to statistically reason and communicate, and the valuable time spent on manipulatives can also sustain long-term effects on building students' confidence in learning statistics and deepening their statistical understanding (Shaw, 2002).

There are fundamental reasons to inherently value the inverted classroom's emphasis on activity-based learning and increased responsibility of the students to become active participants in their own learning. What hasn't been adequately studied is whether and how much the inverted classroom actually has a positive effect on the cognitive and affective outcomes of students.

With the rising enthusiasm for educational reform in statistics education, this dissertation will provide timely insight into the effectiveness of some educational practices in undergraduate statistics education nationwide and identify factors that facilitate or hinder this effectiveness. The intellectual merit of this dissertation is both evident and substantial. Findings from this dissertation can meaningfully inform educators in other disciplines, assisting them in the reform of their own particular conceptualizations and implementations of innovative instructions.

Summary of Principal Findings

Study 1

This study uses a randomized experiment to determine if differences in students' achievement in undergraduate statistics class exist when students learn statistical concepts using virtual manipulatives compared to when students learn statistical concepts using physical manipulatives. The researcher randomly assigned students in different sections to either a physical manipulative condition or a virtual manipulative condition.

The results of this study revealed that there were no significant differences between the BAU group who received traditional concrete manipulatives and the experimental group who received online virtual manipulatives. This study included several student background variables (i.e., gender, age and high school ACT mathematics score) for the examination of interactions with treatment condition. Nonetheless, there were no statistically significant interaction effects between types of manipulatives and any of these background variables. In particular, there was no statistically significant interaction effects between types of manipulatives and high school ACT mathematics scores, informing the literature that ability levels neither intensify nor weaken the effects of types of manipulatives.

The result of no significant difference in GPA one year later refers to the exploration of the long-term effects of types of manipulatives and also performance in that course. The results of the study did not show a significant difference in GPA one year later between the experimental group and the BAU group. The results of the study did demonstrate, nonetheless, that performance (regardless of types of manipulatives) in that course had positive impact on GPA one year later.

Overall, when it comes to manipulatives, their physicality seems unimportant—their manipulability and meaningfulness make them educationally effective supports (Martin, 2009).

Study 2

This research studied use of the flipped classroom in undergraduate statistics class and its effect on student achievement. The results of this study revealed that there were some significant differences between the BAU group in a traditional lecture-type classroom and the experimental group in an inverted classroom. We compared all seven outcomes for the two groups: projects average, tests average, classwork, midterm attendance average, class final attendance average, midterm grade and class final grade. The results for the absolute treatment effects of inverted classroom against traditional classroom show that students in the traditional classroom did better than students in the inverted classroom in projects average, classwork, midterm attendance average, midterm grade and class final grade. We used three different blocks with student background variables as predictors. For the first one, individual student background was explained by age, gender and ethnicity. High school background variables explained by high school GPA and ACT mathematics scores. The third one, university program background, was explained by university cumulative GPA and student major.

After controlling for student background variables of gender, age and ethnicity, students in the traditional classroom did better than students in the inverted classroom in projects average, overall classwork and midterm grade. The model when controlling for student high school background variables of high school GPA and ACT mathematics score, showed that students in the traditional classroom did better than students in the inverted classroom in projects average, overall classwork and midterm grade. Finally, after controlling for student university background variables of student major and university cumulative GPA, students in the traditional classroom performed similarly to students in the inverted classroom in projects average, test average, overall classwork, midterm attendance average, class final attendance, midterm grade and class final grade. When controlling for all (i.e., student background variables, student high school background variables, and university program variables), students in the traditional classroom did better than students in the inverted classroom in midterm grade only.

Practical Implications

Study 1

The foundational position of the study is that when students can visualize a statistical concept in action, a deeper level of understanding occurs. We believe that technology, in the form of virtual manipulatives, may act as an essential component of enhancing statistics instruction by ensuring students' understanding of statistical concepts. This study thus compared the effectiveness of using concrete and virtual manipulatives in undergraduate level statistics class. Virtual manipulatives and traditional manipulatives are equally effective and do not produce long-term differences academically. Still, these results indicate some advantages of the use of virtual manipulatives attractive. Virtual manipulatives can provide feedback to students immediately upon rendering their response. Virtual manipulatives are also dynamic, interactive, flexible and easy to manage. Finally, virtual manipulatives are very affordable, making them a good choice in a budget-tight environment. Overall, the advantages of their use in the classroom are promising in the search for new ways of teaching and learning statistics.

Study 2

Because traditional and inverted classroom are equally effective, educational authorities may not need to promote inverted instruction as a major educational reform. Instead, efforts to promote research and development to engage in the improvement of inverted classroom may become a priority. The teaching assistants who help in the inverted classroom may need training on the tenets of the flipped classroom to create a dynamic, inquiry-based learning environment. The instructor is highly expected to use recitation time for the inverted classroom group as an active session where students would work collaboratively in groups, present statistics problems to the class or in their groups and have active whole-class discussions (not lectures). Instructors may also need training of these inquiry-based and collaborative learning techniques. The instructor needs to be trained as a content expert and a pedagogical expert (and sometimes even a classroom manager) at the same time.

Limitations and Suggestions

Study 1

The purpose of this study was to compare the effectiveness of concrete manipulatives and virtual manipulatives when teaching statistics to undergraduate students in a core content course. The results of the study may not be generalized to the population of all undergraduate students. It also gives no indication of how the results would generalize to other content domains. Further studies may explore along these lines of inquiry regarding the effects of virtual manipulatives in comparison with concrete manipulatives. Research shows that consistent use of manipulatives provides more benefits than temporary use (Sowell, 1989). With only one semester of using virtual manipulatives in this study, the duration may not be classified as consistent use. Further studies may seek some longer period of using virtual manipulatives.

Study 2

The purpose of this study was to compare the effectiveness of traditional lecture-type classroom and inverted classroom when teaching statistics to undergraduate students in a core content course. The results of the study may not be generalized to the population of all undergraduate students and other content domains. Further studies may explore along these lines of inquiry regarding the effects of the traditional classroom in comparison with inverted. The lack of statistically significant advantages of the inverted classroom over the traditional classroom may to a large extent indicate that with only one semester used to compare the two teaching methods in this study, the duration may not be sufficient. Further studies may seek some longer period of using and comparing the two teaching methods. Standardized assessment tools also have a critical role to play in further research on the inverted classroom. Standardized experimental procedures can also be important for future research efforts (e.g., the use of the same teaching assistants for the two different groups). Finally, it is worth emphasizing the largest limitation in the present study—the lack of pretest data. This lack could be a major reason why the advantages of the inverted classroom were unseen in the present study. A pretest-posttest research design is highly desirable to further investigate the concept of the inverted classroom.

Appendix A

Tables with Activities and Learning Outcome Totals

Business as Usual (BAU): Traditional Classroom

24 Lecture times

14 Recitations

3 Review days

2 Test days

Module	**Date**		**Activity of the Day**	**Learning Outcomes**	**Recitations**
Introduction	**August**	28	Opening Day- L1	Module 1-Learning outcomes 1 to 3	
	September	2	Recitation for sections 31 and 35 ONLY during Lecture time		
	September	4	L2	Module 1-Learning outcomes 1 to 3	Recitation1
Human	**September**	9	L3	Module 1-Learning outcomes 1 to 3	
Inference	**September**	11	L4	Module 1-Learning outcomes 4 to 7	Recitation2
	September	16	L 5	Module 1-Learning outcomes 4 to 7	
	September	18	L6	Module 1-Learning outcomes 8 to 10	Recitation3
	September	23	Review Day for Exam 1	Module 1-Learning outcomes	
	September	25	**Exam 1**	Module 1-Learning outcomes	Recitation4
Confidence	**September**	30	L7	Module 2-Learning outcomes 1 to 3	
Intervals	**October**	**2**	**L8**	Module 2-Learning outcomes 1 to 3	Recitation**5**

	October	7	L9	Module 2-Learning outcomes 1 to 3	
	October	9	L10	Module 2-Learning outcomes 4 to 7	Recitation6
	October	14	L11	Module 2- Learning outcomes 4 to 7	
	October	16	L12	Module 2- Learning outcomes 8 to 10	Recitation 7
	October	21	L13	Module 2- Learning outcomes 8 to 10	
	October	23	L14	Module 2- Learning outcomes 8 to 10	Recitation 8
	October	28	Review Day for Exam 2	Module 2-Learning outcomes	
	October	30	**Exam 2**	Module 2- Learning outcomes	Recitation 9
	November	4	L15	Module 3- Learning outcomes 1 to 3	
	November	6	L16	Module 3- Learning outcomes 1 to 3	Recitation 10
	November	11	L17	Module 3- Learning outcomes 1 to 3	
Formal Inference	**November**	13	L18	Module 3- Learning outcomes 1 to 3	Recitation 11
	November	18	L19	Module 3- Learning outcomes 4 to 7	
	November	20	L20	Module 3- Learning outcomes 4 to 7	Recitation 12
	November	25	L21	Module 3- Learning outcomes 4 to 7	
	December	2	L22	Module 3- Learning outcomes 8 to 10	
	December	4	L23	Module 3- Learning outcomes 8 to 10	Recitation 13
	December	9	L24	Module 3- Learning outcomes 8 to 10	
	December	11	Review	Module 3- Learning outcomes	Recitation 14
	December	15	**FINAL Exam 3 from 10:30-12:30**	Module 1,2, 3- Learning outcomes	

Experimental (EXP): Inverted Classroom

24 Lecture times

14 Recitations

3 Review days

2 Test days

Module	Date		Activity of the Day	Learning Outcomes	Recitations
Introduction	**August**	27	Opening Day- L1	Module 1-Learning outcomes 1 to 3	
	September	3	L2	Module 1-Learning outcomes 1 to 3	Recitation 1
	September	8	L3	Module 1-Learning outcomes 1 to 3	
	September	10	L4	Module 1-Learning outcomes 4 to 7	Recitation 2
	September	15	L 5	Module 1-Learning outcomes 4 to 7	
	September	17	L6	Module 1-Learning outcomes 8 to 10	Recitation 3
	September	22	Review Day for Exam 1	Module 1-Learning outcomes	
	September	24	**Exam 1**	Module 1-Learning outcomes	Recitation 4
	September	29	L7	Module 2-Learning outcomes 1 to 3	
	October	**1**	**L8**	Module 2-Learning outcomes 1 to 3	**R**ecitation **5**
	October	6	L9	Module 2-Learning outcomes 1 to 3	
	October	8	L10	Module 2-Learning outcomes 4 to 7	Recitation 6
Confidence Intervals	**October**	13	L11	Module 2- Learning outcomes 4 to 7	
	October	15	L12	Module 2- Learning outcomes 8 to 10	Recitation 7
	October	20	L13	Module 2- Learning outcomes 8 to 10	
	October	22	L14	Module 2- Learning outcomes 8 to 10	Recitation 8
	October	27	Review Day for Exam 2	Module 2-Learning outcomes	
	October	29	**Exam 2**	Module 2- Learning outcomes	Recitation 9

Formal Inference	**November**	3	L15	Module 3- Learning outcomes 1 to 3	
	November	5	L16	Module 3- Learning outcomes 1 to 3	Recitation 10
	November	10	L17	Module 3- Learning outcomes 1 to 3	
	November	12	L18	Module 3- Learning outcomes 1 to 3	Recitation 11
	November	17	L19	Module 3- Learning outcomes 4 to 7	
	November	19	L20	Module 3- Learning outcomes 4 to 7	Recitation 12
	November	24	L21	Module 3- Learning outcomes 4 to 7	
	December	1	L22	Module 3- Learning outcomes 8 to 10	
	December	3	L23	Module 3- Learning outcomes 8 to 10	Recitation 13
	December	8	L24	Module 3- Learning outcomes 8 to 10	
	December	10	Review	Module 3- Learning outcomes	Recitation 14
	December	15	**FINAL Exam 3 from 3:30-5:30**	Module 1,2, 3- Learning outcomes	

Appendix B

Description of Modules and Learning Outcomes

Module 1 – Human Inference

Overarching Goal

The primary intent of this module is to develop the skills needed to absorb common statistical information and to correctly form the associated human inferences.

Learning Outcomes

You will know you have successfully completed this module when you are able to:

1. Identify categorically good or bad statistical summaries, charts and graphs and explain the reasons they are so categorized.
2. Identify categorically good or bad statistical arguments based on statistical summaries, charts, and graphs, and explain the reasons they are so categorized.
3. Compute basic statistical summaries and create simple graphs.
4. Define and apply basic experimental design vocabulary.
5. Identify confounding variables and evaluate their effects on experimental results.
6. Explain the role of randomization in simple experimental design.
7. Explain in non-mathematical terms the concept of statistical significance.
8. Identify and assess associations seen in scatterplots and two-way tables.
9. Distinguish the concepts of association and causation, and explain how they offer different types of evidence.
10. Compute, apply, and interpret the correlation coefficient.

Duration – Minimum of 4 weeks

Module 2 – Confidence Intervals

Overarching Goal

The primary intent of this module is to develop a broad understanding of what statistical confidence means, what it doesn't mean, and what components are required for its construction.

Learning Outcomes

You will know you have successfully completed this module when you are able to:

1. Define and demonstrate simple random sampling.
2. Identify and analyze alternative sampling methods.
3. Explain the difference between randomness and representativeness.
4. Define sampling variability and explain the role it plays in the construction of a confidence interval.
5. Define sampling distribution and explain the role it plays in the construction of the margin of error.
6. Compute and interpret confidence intervals for a proportion or mean.
7. Define and apply the empirical rule to solve probability problems.
8. Identify categorically good or bad surveys and explain the reasons they are so categorized.
9. Explain the difference between sampling variability and non-sampling variability.
10. Identify and evaluate strategies for addressing non-sampling variability.

Duration: Minimum of 4 weeks

Module 3 – Formal Inference

Overarching Goal

The primary intent of this module is to understand the conceptual tenets and practical consumption of statistical hypothesis testing, beginning with more accessible concepts of sensitivity and specificity.

Learning Outcomes

You will know you have successfully completed this module when you are able to:

1. Define and compute sensitivity and specificity.

2. Explain the effect on sensitivity and specificity of changes to the testing criteria.
3. Identify and demonstrate the difference between probabilities of conditional and unconditional events.
4. Define Type I error and explain how to view hypothesis testing as a screen test.
5. Explain the difference between a Type I error and a p-value.
6. Define the meaning of the phrase statistical significance.
7. Analyze the use of the phrase "statistically significant" in media reports.
8. Explain the difference between statistical significance and practical significance.
9. Execute the steps needed to test simple hypothesis.
10. Compute and demonstrate the use of p-value when testing a hypothesis.

Duration: Minimum of 4 weeks

Appendix C

Videos and URL's for them

Video	YouTube URL
1. Number Sense – Basic Numeracy	https://www.youtube.com/watch?v=_kVvYTYkEWY
2. Number Sense – Computations and Benchmarks	https://www.youtube.com/watch?v=aHBhDx_Potk
3. Confounding and the Language of Experiments – Introduction	https://www.youtube.com/watch?v=xV2bln1BPbw
4. Confounding and the Language of Experiments – Comparison and Randomization	https://www.youtube.com/watch?v=eR29xVBZU1E
5. Confounding and the Language of Experiments – Statistical Significance	https://www.youtube.com/watch?v=xgmXDHfqXVQ
6. Correlation and Causation - Scatterplots	https://www.youtube.com/watch?v=FDcM7wcCd7E
7. Correlation and Causation – Correlation Coefficient	https://www.youtube.com/watch?v=qRpr-8krjVU
8. Correlation and Causation - Causation	https://www.youtube.com/watch?v=njqhb2KsOsg
9. Sampling Content - Introduction	https://www.youtube.com/watch?v=rJGDr0cM9VM
10. Sampling Content – Language and Techniques	https://www.youtube.com/watch?v=QNCxfDxDfbs
11. Sampling Content – Confidence Intervals	https://www.youtube.com/watch?v=9dbttMpR4KM
12. Sampling Content – When the MOE Doesn't Apply	https://www.youtube.com/watch?v=Y9gbUoK8teA
13. Sensitivity and Specificity – Introduction and Definitions	https://www.youtube.com/watch?v=VgGZksveZaE
14. Sensitivity and Specificity – Computations and Examples	https://www.youtube.com/watch?v=otmwzs1HhyQ
15. Hypothesis Testing – As a Diagnostic Test	https://www.youtube.com/watch?v=KATz04jrsOk
16. Hypothesis Testing – Applying the Concepts	https://www.youtube.com/watch?v=g8VYhcvqg9o

17. Hypothesis Testing – Practical Significance	https://www.youtube.com/watch?v=RL4j7qDv748
18. Hypothesis Testing – Computations	https://www.youtube.com/watch?v=41Yw8i9szEQ

Appendix D

Description of Activities applied during Main-Class time and Recitation time

	BAU Treatment	EXP Treatment
Time Allocated for Main-Class Content	A. Course content videos are shown in class at times appropriate for the material. This amounts to roughly one content video every two main-class days. Content video will take approximately 15 minutes. B. Daily explanations and examples prepared and delivered by PowerPoint. These are to be 10 minutes in length. Same as B. in EXP. C. Questions allowed from students for approximately 10 minutes. Same as C in EXP. Time: approximately 35 minutes on content video days, and 20 minutes on other days.	A. Course content videos are assigned to be watched outside of class at times appropriate for the material. This amounts to roughly one content video every week. **Requires no class time.** B. Daily explanations and examples prepared and delivered by PowerPoint. These are to be 10 minutes in length. Same as B. in BAU. C. Questions allowed from students for approximately 10 minutes. Same as C in BAU. Time: approximately 20 minutes each day.
Remaining Main-Class Time	A. The instructor will use PowerPoint (or a document projector) to go over Beyond the Numbers assignments that students	A. Students may be allowed to complete BN(s) in class in small groups, or through

	have performed as homework or completed as small groups in the recitations and submitted to CPR. This must be a static presentation. B. Students may be allowed to work alone on BNs in class (no collaboration). C. Instructor and TAs generally should avoid facilitating active engagement with BNs in class. D. A select group of BNs* must not be facilitated or discussed in class. Time: approximately 15 minutes on content video days and 30 minutes on other days.	some interactive structure, and submitted to CPR after class. B. Students may be allowed to complete BN(s) before class and have them submitted to CPR. In this case, the instructor and TAs can use an active-learning activity to surface the import of those completed activities. C. Instructor and TAs generally should facilitate active engagement with BNs in class. D. A select group of BNs* must be facilitated and discussed in class. Time: approximately 30 minutes.
	*BNs: 1.9, 1.21, 1.24; 2.7, 2.18, 2.19; 3.8, 3.24, 3.26	

	BAU Treatment	EXP Treatment
Recitation Activities	A. TAs generally should facilitate active engagement with BNs in recitation. B. A select group of BNs* must be facilitated and discussed by the TA in recitation. C. The TA is allowed to answer questions about CPR, Blackboard, or illustrate computations. D. TA may be allowed to supervise student completion of selected BN(s) in recitation. E. TA will distribute weekly quizzes created by the instructor. Time: approximately 50 minutes.	A. TAs generally should not facilitate active engagement with BNs in recitation. B. A select group of BNs* must not be facilitated or discussed by the TA in recitation. C. The TA is allowed to answer questions about CPR, Blackboard, or illustrate computations. D. TA may be allowed to supervise student completion of selected BN(s) in recitation. E. TA will distribute weekly quizzes created by the instructor. Time: approximately 50 minutes.
	*BNs: 1.9, 1.21, 1.24; 2.7, 2.18, 2.19; 3.8, 3.24, 3.26	

Beyond the Numbers Referenced above:

1.9 Why Numeracy Matters

1.21 Random Reflections

1.24 What to Believe

2.7 Are Online Reviews Statistical Samples?

2.18 Mathematically Organic Bells

2.19 Confidence in Repetition

3.8 Thinking about Conditional Reasoning

3.24 Accept or Fail to Reject? Semantics or Real?

3.26 Error Rates and P-Values

Appendix E

Day-by-Day Explanation of BAU and EXP

BAU			EXP		
Activity of the Day BAU	**Homework BAU**	**Recitation BAU**	**Activity of the Day EXP**	**Homework EXP**	**Recitation EXP**
Opening Day- L1 • Syllabus • Introducing the team • Why is all this important? Two videos from YouTube (Fun with Math and “did you know” with mistake)	BN 1.2		Opening Day- L1 • Syllabus • Introducing the team Why is all this important? Two videos from You Tube(Fun with math and “did you know” with mistake)	BN 1.1 video Watch and complete all questions	
L2 • Show video BN 1.1 • Instructor shows PPT 2 (Material in video 1+some vocabulary) all this will help students complete BN 1.1 as notes in class • Instructor goes over BN 1.2 using document camera.	BN1.3 BN1.1	R1 **BN 1.9** BN 1.17 Discovery type of activity lead by TA	L2 • Instructor shows PPT 2 covering material in video BN 1.1 • Students work in groups BN 1.2 • Students present and discuss problems • Students work in groups BN 1.3 • Students present and discuss problems	BN 1.4 video Watch and complete all questions; BN 1.5	R1

L3 • Show video BN1.4 • Instructor shows PPT 3(material covered in the video); all this will help students complete BN 1.4 as notes in class. • Instructor goes over BN 1.3 using document camera • Instructor goes over BN 1.5 using document camera	BN 1.6 BN1.8 BN 1.4		L3 • Instructor -PPT 3 covering material in BN 1.4 video and questions from the BN 1.4 • Students work in groups BN 1.6 • Students work in groups **BN 1.9 discovery type of activity lead by Instructor** • Class discussion over questions from BN 1.6 and BN 1.9	BN 1.13video Watch and complete	
L4 • Show video BN 1.13 • Show video BN 1.16 • Instructor explains PPT 4 material covered in the two videos above and will help students complete notes BN 1.13 and BN 1.16 • BN 1.14 individually or instructor demonstrates using document camera	BN 1.13 BN 1.16 maybe complete parts of BN 1.14 all the whole assignment	R2 **BN 1.21** **Discovery type of activity lead by TA** **BN 1.24** Handout1	L4 • Instructor -PPT 4 covering material in video BN 1.13 and questions in BN 1.13 • Students work in pairs BN 1.14 • Students report answers and discuss different responses • Handout1-group work	BN 1.16 Watch and complete	R2 BN 1.26 BN 1.29

			• Students report answers to class		
L 5 • Instructor-Show video BN 1.19 • Instructor-Show video BN 1.25 • PPT 5 material covered in BN 1.19 and BN 1.25 videos above and will help students complete the notes. • Instructor demonstrates one problem from BN 1.20 using the PPT presentation	BN 1.20 BN 1.19 BN 1.25		L5 • Instructor-PPT 5 covering material in video BN 1.16 and questions in BN 1.16 • BN 1.17 students work in groups • Handout2-group work • Students report results in front of the class and open a discussion.	BN 1.19 and BN1.25 Watch and complete BN 1.20	
L6 • Show video BN 1.28 • Show video 1.31 • PPT 6 material covered in BN 1.28 and BN 1. 31 video and will help students to complete BN's	BN 1.26 BN 1.29 BN 1-28 BN 1-31	R3 **BN 1.24 Discovery type of activity lead by TA**	L 6 • Instructor-PPT 6 covering material in BN 1-19 and BN 1-25 and questions in the two BN's • Class works in groups or individually **BN 1.21** • Students work in groups- **BN 1.24 discovery type of activity lead by instructor**	BN 1.28 and BN1.31 videos Watch and complete Sympson's paradox PPT	R3 BN 1.30 BN 1.32 BN 1.33

			• Class discussion going over selected problems from BN's completed in class		
Review Day for Exam 1 • Instructor-Go over different examples posted as a review using document camera • May use BN 1.30; BN 1.32 or BN 1.33 to demonstrate in class using a document camera			Review for Test 1 • Simpson's paradox PPT-instructor explains • BN 1.34-instructor works with students and completes as a discussion • Review Test 1-students work in groups and later present on the board	BN 1.35 HW	
Exam 1 –Same as EXP		R4 Simpsons paradox PPT presented from TA's BN 1.34 BN 1.35	Exam 1-same as BAU	Watch and complete BN 2.1	R4
L7 • Results from Test 1 • Show video BN 2.1	BN 2.3 BN 2.5 BN 2.1		L7 • Test 1 results	BN 2.5 video	

• PPT 7 material in video BN 2.1 and will help students complete notes BN 2.1 • Instructor demonstrates BN 2.2 using document camera				• Instructor-PPT 7 covering material in video BN 2.1 • Students work together in groups BN 2.2 • Students work together in groups BN 2.3 • Class was split in advance and each group reports a specific question. Instructor leads a discussion.	Watch and complete BN 2.4	
L8 • Show video BN 2.4 • PPT 8 covers material in video above and helps students complete the BN 2.4 • Practice Handout –Instructor demonstrates using document camera	BN 2.6	R5 BN 2.4 **BN 2.7** **Discovery type of activity lead by TA**	L8 • PPT 8 covering material in video BN 2.4 • Go over BN 2.5 with students. They share answers and instructor facilitates discussion • Work in groups-BN 2.6 • Students report answers on specific exhibits.	BN 2.7	R5	
L9 • PPT 9-Instructor • Handout –Instructor first allows students to work individually	Think about the project		L9 • Instructor shows and explains PPT 9	Watch video and complete BN 2.13		

and then demonstrates answers for them to check using document camera			• **Students work with instructor on BN 2.7** • Handout3-work in groups • Students report selected problems and lead discussion		
L10 • Show video BN 2.13 • PPT covers material in video BN2.13 • Instructor explains BN 2.14	BN 2.15 BN 2.13	R6 **BN 2.18 Discovery type of activity lead by TA** BN 2.16	L10 • PPT 10 covering material in video BN 2.13 • BN 2.14 • BN 2.15	BN 2.16	R6
L11 • Instructor shows BN 2.21 using a document camera • PPT 11-Instructor summarizes material	BN 2.22		L11 • **Instructor leads BN 2.18** • Students work in groups **BN 2.19** • Students report answers and instructor open a discussion	Watch video and complete BN 2.25	

L12 • Show video BN 2.25 • PPT 12 covers material in BN 2.25 will help students to complete BN 2.25 as notes. • BN 2.26 instructor demonstrates or students individually	BN 2.25	R7 **BN 2.19 Discovery type of activity lead by TA**	L12 • BN 2.25 go over student presentation • BN 2.26	BN 2.21 BN 2.22	R7
L13 • Instructor-PPT 13 • Instructor demonstrates using document camera BN 2.27 • BN 2.28 Instructor/Individual	BN 2.28 Work on project		L13 • Half class-Group work BN 2.27 • Other half of class works in groups-BN 2.28 • Students report selected problems by groups	Work on project	
L14 • Instructor-PPT 14 • Instructor demonstrates using document camera BN 2.30 • BN 2.31 Instructor/Individual	Work on project	R8	L14 • Students work in pairs.BN 2.30 • Students work in pairs BN 2.31 • Instructor leads class discussion, going over problems with students		R8
Review Day for Exam 2			Review Day for Exam 2		

• Instructor goes over selected problems using document camera			• Students are allowed to work individually or in groups. Later they present selected problems on the board.		
Exam 2		R9	Exam 2	Watch video and complete BN 3.1	R9
L15 • Test 2 results • Show video BN 3.1 • PPT 15 covering material in Video BN 3.1 and questions in the same BN • BN 3.2 Instructor demonstrates using document camera	BN 3.1		L15 • Test 2 results • PPT covering video BN 3.1 • Students work in groups BN 3.2 • Students share results and conclude with a discussion and questions	Watch video and complete BN 3.3	
L16 • Show video BN 3.3 • PPT 16 covering material in BN 3.3 video and questions in the same BN	BN 3.3	R10 BN 3.7 BN 3.4 BN 3.5	L16 • PPT covering BN 3.3 • Students work with instructor on BN 3.4 • Students report answers for BN 3.4	BN 3.4 complete with data	R 10

• BN 3.6 Instructor completes and demonstrates using document camera			• Students work in groups BN 3.5 • Students report answers in front of the class		
L17 • Show BN 3.13 video • PPT 17 covering material in BN 3.13 video and questions in the same BN • BN 3.14-demonstration from instructor using document camera	BN 3.15 BN 3.13		L17 • Students work in pairs BN 3.6 • Students work with instructor **BN 3.8** • Discussion and result reports from students	BN 3.9	
L18 • Show video BN 3.16 • PPT 18 covering video BN 3.16 and questions in the same BN • BN 3.17-instructor reviews with students, using a document camera	BN 3.18 BN 3.16	R11 **BN 3.8** **Discovery type of activity lead by TA** BN 3.9	L18 • Class completes in groups BN 3.7 • Groups report results	Watch video and complete BN 3.13	R11
L19 • Instructor summarizes material PPT 19 • Instructor demonstrates BN 3.20 using document camera	BN 3.19		L19 • Instructor-PPT based on video BN 3.13 • Students work together BN 3.14	BN 3.16 Watch video and complete	

• Individual work BN 3.19 • Instructor demonstrates problems from BN 3.19			• Students work in pairs BN 3.15 • Students demonstrate answers on the board and open discussion		
L20 • Show BN 3.21 video • PPT 20 covering material in the video BN 3.21 and questions in the same BN • Instructor demonstrates BN 3.22 using document camera	BN 3.23	R12	L20 • PPT based on video BN 3.16 • Students work in groups BN 3.17 • Students work in groups BN 3.18 • Instructor leads discussion, and students present responses to different exhibits.	BN 3.19 Watch video and complete BN 3.21	R12
L21 • Students work individually Handout4 • Instructor reviews selected problems using document camera			L21 • PPT based on video BN 3.21 • Students work in pairs BN 3.20 • Students work in groups BN 3.22 • Students report results	BN 3.23 Watch video and complete BN 3.27	
L22 • Show video BN 3.27	BN 3.27 BN 3.29		L22	BN 3.29	

• PPT 22 covering video BN 3.27 material and questions • Instructor goes over problems in BN 3.28 using document camera			• PPT based on video BN 3.27 • Students work in groups BN 3.28 • **Instructor leads BN 3.24** • Groups report		
L23 • Instructor demonstrates BN 3.29 using document camera • Individual work BN 3.30		R13 **BN 3.24** **Discovery type of activity lead by TA**	L23 • Check together BN 3.29 • **Instructor leads BN 3.26** • Students work in groups BN 3.30 • Students report results	Work on project	R13
L24 • Instructor demonstrates BN 3.31 using document camera • BN 3.32 Individual work	BN 3.34		L24 • Half of the class - Students work in small groups a specific problem from BN 3.31 • Other half of the class- Students work in small groups BN 3.32 • Each group reports a problem on the board	BN 3.34	

			• discussion		
Review for test 3 • Instructor demonstrates selected problems using document camera		R14 **BN 3.26** **Discovery type of activity lead by TA**	Review for test 3 • Students work on different problems and then present answers on board and discuss different solutions and responses.		R14
FINAL Exam 3 from 10:30-12:30			FINAL Exam 3 from 3:30-5:30		

Appendix F

Technology Use in Both Class Settings

Technology use	Inverted	Traditional
Videos for presenting subject material	Yes	Yes
PowerPoint presentations for subject material-10 min. – Instructor will present in both classes	Yes	Yes
Videos used to show applications, but they are supplement material – such as the “Split Brain” example	Yes	Yes
CPR (Peer Review System/Computer based) projects – at least two And Beyond the Numbers Homework	Yes	Yes
Blackboard-based quizzes (online)	Yes	Yes
Using Safe Assign for turning in projects – at least two	Yes	Yes
Web quests and research for projects –at least two	Yes	Yes
“Survey Monkey” web application for creating surveys	Yes	Yes
Excel or other graph-making software for organizing data	Yes	Yes
Website-based application such as the one for simulating confidence intervals: http://statweb.calpoly.edu/chance/applets/Confsim/Confsim.html	Yes	Yes
Blackboard Discussion board	Yes	Yes
Blackboard announcements and additional materials needed for specific assignments	Yes	Yes

Appendix G-Complete List of Items used to measure competence in both groups

MODULE 1	MODULE 2	MODULE 3
Beyond the Numbers 1.1	Beyond the Numbers 2.1	Beyond the Numbers 3.1
Beyond the Numbers 1.2	Beyond the Numbers 2.2	Beyond the Numbers 3.2
Beyond the Numbers 1.3	Beyond the Numbers 2.3	Beyond the Numbers 3.3
Beyond the Numbers 1.4	Beyond the Numbers 2.4	Beyond the Numbers 3.4
Beyond the Numbers 1.5	Beyond the Numbers 2.5	Beyond the Numbers 3.5
Beyond the Numbers 1.6	Beyond the Numbers 2.6	Beyond the Numbers 3.6
Beyond the Numbers 1.8	Beyond the Numbers 2.7	Beyond the Numbers 3.7
Beyond the Numbers 1.9	Beyond the Numbers 2.13	Beyond the Numbers 3.8
Beyond the Numbers 1.13	Beyond the Numbers 2.14	Beyond the Numbers 3.9
Beyond the Numbers 1.14	Beyond the Numbers 2.15	Beyond the Numbers 3.13
Beyond the Numbers 1.16	Beyond the Numbers 2.16	Beyond the Numbers 3.14
Beyond the Numbers 1.17	Beyond the Numbers 2.18	Beyond the Numbers 3.15
Beyond the Numbers 1.19	Beyond the Numbers 2.19	Beyond the Numbers 3.16
Beyond the Numbers 1.20	Beyond the Numbers 2.21	Beyond the Numbers 3.17
Beyond the Numbers 1.21	Beyond the Numbers 2.22	Beyond the Numbers 3.18
Beyond the Numbers 1.24	Beyond the Numbers 2.25	Beyond the Numbers 3.19
Beyond the Numbers 1.25	Beyond the Numbers 2.26	Beyond the Numbers 3.20
Beyond the Numbers 1.26	Beyond the Numbers 2.27	Beyond the Numbers 3.21
Beyond the Numbers 1.28	Beyond the Numbers 2.28	Beyond the Numbers 3.22
Beyond the Numbers 1.29	Beyond the Numbers 2.30	Beyond the Numbers 3.23
Beyond the Numbers 1.30	Beyond the Numbers 2.31	Beyond the Numbers 3.24

Beyond the Numbers 1.31		Beyond the Numbers 3.26
Beyond the Numbers 1.32		Beyond the Numbers 3.27
Beyond the Numbers 1.33		Beyond the Numbers 3.28
Beyond the Numbers 1.34		Beyond the Numbers 3.29
Beyond the Numbers 1.35		Beyond the Numbers 3.30
		Beyond the Numbers 3.31
		Beyond the Numbers 3.32
		Beyond the Numbers 3.34

Appendix H

Survey of Attitude toward Statistics

Thinking about your views on statistics, to what extent do you agree with the following statements? (Please check only one box in each row.)	Strongly Agree	Agree	Disagree	Strongly Disagree
I enjoy reading about statistics.				
Making an effort in statistics is worth it because it will help me in the work that I want to do later on.				
I look forward to my statistics lessons.				
I do statistics because I enjoy it.				
Learning statistics is worthwhile for me because it will improve my career prospects.				
I am interested in the things I learn in statistics.				
Statistics is an important subject for me because I need it for what I want to study later on.				
I learn many things in statistics that will help me get a job.				
Thinking about studying statistics, to what extend do you agree with the following statements? (Please check only one box in each row.)	Strongly Agree	Agree	Disagree	Strongly Disagree
I often worry that statistics classes will be difficult for me.				
I am just not good at statistics.				
I get very tense when I have to do statistics assignments.				
I get good grades in statistics.				

I get very nervous doing statistics problems.				
I learn statistics quickly.				
I have always believed that statistics is one of my best subjects.				
I feel helpless when doing a statistics problem.				
In my statistics class, I understand even the most difficult work.				
I worry I will get poor grades in statistics.				

Thinking about your experience in this statistics course (SAT 210), to what extend do you agree with the following statements? (Please check only one box in each row.)	Strongly Agree	Agree	Disagree	Strongly Disagree
This way of learning statistics prompted me to ask questions in class.				
This way of learning statistics motivated me to express my views or opinions in class.				
This way of learning statistics made me want to interact with the instructor.				
This way of learning statistics made me feel that statistics makes sense to me now.				
This way of learning statistics made me feel bored in class.				
This way of learning statistics was strange to me.				
I often came to class without completing readings or assignments.				
I wish all statistics courses could be offered in this way.				

REFERENCES

Ajzen, I. (1989). Attitude structure and behavior. A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), Attitude Structure and Function (pp. 241-274*). Hillsdale, NJ: Lawrence Earlbaum.*

Allen, I. E., & Seaman, J. (2007). Making the grade: Online education in the United States, *2006*. *Sloan Consortium. PO Box 1238, Newburyport, MA 01950.*

Ambrose, S. A., Bridges, M. W., DiPietro, M., Lovett, M. C., Norman, M. K., Mayer, R. E. (2010) How Learning Works: Seven Research-Based Principles for Smart Teaching, *Jossey-Bass.*

Anderson, L. W. (2005). Objectives, evaluation, and the improvement of education. *Studies in educational evaluation, 31(2-3), 102-113.*

Araki, L. R., & Schultz, K. S. (1995, April). Students' attitudes towards statistics and their retention of statistical concepts. *Paper presented at the annual meeting of the Western Psychological Association, Los Angeles, CA.*

Arnold-Garza. 2014. "The Flipped Classroom." *College & Research Libraries News 75, no. 1: 10-13.*

Baker, J. W. (2000). The 'classroom flip': Using web course management tools to become a guide by the side. *Paper presented at the 11th international conference on college teaching and learning, Jacksonville, FL.*

Baker, B. D., Keller-Wolff, C., & Wolf-Wendel, L. (2000). Two steps forward, one step back: Race/ethnicity and student achievement in education policy research. *Educational Policy*, *14*(4), 511-529.

Batanero, C., Burrill, G., & Reading, C. (Eds.). (2011). Teaching statistics in school mathematics-challenges for teaching and teacher education: A joint ICMI/IASE study*: the 18th ICMI study* (Vol. 14). *Springer Science & Business Media.*

Baxter Hastings, N., Gordon, F. S., Gordon, S. P., Narayan, J., & Mathematical Association of America. (2006). A fresh start for collegiate mathematics: Rethinking the courses below calculus. *Washington, DC: Mathematical Association of America*

Bennet, J. (2011, August 29). Work experience is what international students want. *Campus Review*. http://www.campusreview.com.au/blog/news/work-experience-is-what-international-students-want/

Bergmann, J., Overmyer, J., & Wilie, B. (2013). The flipped class: What it is and what it is not. *The Daily Riff, 9.*

Bergmann, J., & Sams, A. (2012a). Flip your classroom: Reach every student in every class every day, *International Society for Technology in Education, Washington, D.C*

Berk, R. A., & Nanda, J. P. (1998). Effects of jocular instructional methods on attitudes, anxiety, and achievement in statistics courses.

Berrett, D. 2012. How 'Flipping' the Classroom Can Improve the Traditional Lecture. *Education Digest* 78(1): 36-41.

Bishop, J. L., & Verleger, M. A. (2013, June). The flipped classroom: A survey of the research. *ASEE National Conference Proceedings, Atlanta, GA* (Vol. 30, No. 9, pp. 1-18).

Björklund, C. (2014). Powerful teaching in preschool–a study of goal-oriented activities for conceptual learning. *International Journal of Early Years Education, 22*(4), 380-394.

Bland, D. C. (2006). Researching educational disadvantage: Using participatory research to engage marginalised students with education *(Doctoral dissertation, Queensland University of Technology).*

Boggan, M., Harper, S., & Whitmire, A. (2010). Using manipulatives to teach elementary mathematics. *Journal of Instructional Pedagogies*, *3*, 1.

Bradstreet, T. (1996), "Teaching Introductory Statistics Courses So That Nonstatisticians Experience Statistical Reasoning," *American Statistician*, 50(1), 69-78.

Brakke, K., Wilson, J.H. &Bradley, D.V. (2007). Beyond basics: Enhancing undergraduate statistics instruction. *Best Practices for Teaching Statistics and Research Methods in the Behavioral Sciences*, 109-122

Brame, C. J. (2013). Flipping the classroom. *Retrieved*, *August*, *29*, 2013.

Burns, B. A., & Hamm, E. M. (2011). A comparison of concrete and virtual manipulative use in third-and fourth-grade mathematics. *School Science and Mathematics*, *111*(6), 256-261.

Byrne, G. (1989). US students flunk math, science. *Science*, *243*(4892), 729-730.

Carbonneau, K. J., Marley, S. C., & Selig, J. P. (2013). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *Journal of Educational Psychology*, *105*(2), 380.

Cashin, S. E., & Elmore, P. B. (2005). The Survey of Attitudes Toward Statistics scale: A construct validity study. *Educational and Psychological Measurement*, *65*(3), 509-524.

Castro, A. M. (2006). Preparing elementary preservice teachers to use mathematics curriculum materials. *The Mathematics Educator*, *16*(2).

Chance, B. L. (2005). Integrating pedagogies to teach statistics. *MAA NOTES*, *65*, 101.

Chance, B. L., & Garfield, J. B. (2002). New approaches to gathering data on student learning for research in statistics education. *Statistics Education Research Journal*, *1*(2), 38-41.

Chance, B., Ben-Zvi, D., Garfield, J., & Medina, E. (2007). The Role of Technology in Improving Student Learning of Statistics.

Cheng, Y. C., & Mok, M. M. (2008). What effective classroom? Towards a paradigm shift. *School Effectiveness and School Improvement*, *19*(4), 365-385.

Clements, D. H. (1999). The future of educational computing research: The case of computer programming. *Information Technology in Childhood Education Annual*, *1999*(1), 147-179.

Clements, D. H. (2000). From exercises and tasks to problems and projects: Unique contributions of computers to innovative mathematics education. *The Journal of Mathematical Behavior*, *19*(1), 9-47.

Clements, D. H., & McMillen, S. (1996). Rethinking" concrete" manipulatives. *Teaching Children Mathematics*, *2*(5), 270-279.

Common Core State Standards Initiative (CCSSM) (2010). Common Core State Standards for Mathematics. Washington, DC: National Governors Association for Best Practices and the Council of Chief State School Officers. Available at: http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf

Common Core State Standards Initiative. (2010). Common core state standards for mathematics. Retrieved from http://corestandards.org

Cobb, P. (1991). Reconstructing Elementary School Mathematics. *Focus on Learning Problems in Mathematics*, *13*(2), 3-32.

Cobb, G. (1992). Teaching statistics. *Heeding the Call for Change: Suggestions for Curricular Action*, *22*, 3-43.

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly*, *104*(9), 801-823.

Cohen, M. (1988). Restructuring the Education System: Agenda for the 1990s. *Publication Sales, National Governors' Association, 444 North Capitol Street, Washington, DC 20001-1572*.

Cohen, K. C. (Ed.). (2012). Internet links for science education: Student-scientist partnerships (Vol. 4). *Springer Science & Business Media*.

Crompton, H. (2011). Mathematics in the age of technology: There is a place for technology in the mathematics classroom. *Journal of the research center for educational technology*, *7*(1), 54-66.

D'Andrea, L., & Waters, C. (2002, July). Teaching statistics using short stories: Reducing anxiety and changing attitudes. In *Sixth International Conference on Teaching Statistics, Cape Town, South Africa*.

Dani, B. Z., & Joan, G. (2004). Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges. *The challenge of developing statistical literacy, reasoning and thinking* (pp. 3-15). Springer, Dordrecht.

Day, J. A., & Foley, J. D. (2006). Evaluating a web lecture intervention in a human–computer interaction course. *IEEE Transactions on Education*, *49*(4), 420-431.

Davis, M. (2003). Barriers to reflective practice: The changing nature of higher education. *Active Learning in Higher Education*, *4*(3), 243-255.

Della Cava, M. R. (2012). Sal Khan's 'Academy' sparks a global tech revolution in education. USATODAY.com. Retrieved May 13, 2015, from: URL:

http://www.usatoday.com/life/people/story/2012-05-30/sal-khanprofile-khan-academy/55270348/1

De Veaux, R. D., & Velleman, P. F. (2008). Math is music; statistics is literature (or, why are there no six-year-old novelists?). *Amstat News*, *375*, 54-58.

DelMas, R. C. (2002). Statistical literacy, reasoning, and learning: A commentary. *Journal of Statistics Education*, *10*(3).

Dorward, J. (2002). Intuition and research: Are they compatible? *Teaching Children Mathematics*, *8*(6), 329.

Driscoll, M. (1983). Research within reach: Secondary school mathematics. *Reston, VA: National Council of Teachers of Mathematics*

Durmus, S., & Karakirik, E. (2006). Virtual manipulatives in mathematics education: A theoretical framework. *TOJET: The Turkish Online Journal of Educational Technology*, *5*(1).

Dupuis, D. N., Medhanie, A., Harwell, M., Lebeau, B., Monson, D., & Post, T. R. (2012). A Multi-Institutional Study of the Relationship between High School Mathematics Achievement and Performance in Introductory College Statistics. *Statistics Education Research Journal*, *11*(1).

Emmioğlu, E. S. M. A., & Capa-Aydin, Y. (2012). Attitudes and achievement in statistics: A meta-analysis study. *Statistics Education Research Journal*, *11*(2).

Falk, R., & Greenbaum, C. W. (1995). Significance tests die hard: The amazing persistence of a probabilistic misconception. *Theory & Psychology*, *5*(1), 75-98.

Fink Z (2011) Big thinkers: Salman Khan on liberating the classroom for creativity. Edutopia: K-12 education & learning innovations with proven strategies that work. *Retrieved 9 January 2016. http://www.edutopia.org/salman-khan-academy-flipped-classroom-video*

Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education*, *28*, 96-105.

Forster, P. A. (2006). Assessing technology-based approaches for teaching and learning mathematics. *International Journal of Mathematical Education in Science and Technology*, *37*(2), 145-164.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). Guidelines for assessment and instruction in statistics education (GAISE) report. *Alexandria: American Statistical Association*.

Frederickson, N., Reed, P., & Clifford, V. (2005). Evaluating web-supported learning versus lecture-based teaching: Quantitative and qualitative perspectives. *Higher Education, 50*, 645–664.

Freer Weiss, D. M. (2006). Keeping it real: The rationale for using manipulatives in the middle grades. *Mathematics Teaching in the Middle School, 11*(5), 238-242.

Fulton, K. P. (2012). 10 reasons to flip. *Phi Delta Kappan*, *94*(2), 20-24.

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, *70*(1), 1-25.

Gal, I. (2005). Towards "probability literacy" for all citizens: Building blocks and instructional dilemmas. *Exploring Probability in School* (pp. 39-63). Springer, Boston, MA.

Gal, I., Ginsburg, L., & Schau, C. (1997). Monitoring attitudes and beliefs in statistics education. *The assessment challenge in statistics education*, *12*, 37-51.

Garfield, J. (2003). Assessing Statistical Reasoning. *Statistics Education Research Journal*, 2(1), 22-38

Garfield, J., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education*, 44-63.

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.

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Springer Science & Business Media.

Garfield, J., & Ben-Zvi, D. (2009). Helping students develop statistical reasoning: Implementing a statistical reasoning learning environment. *Teaching Statistics*, *31*(3), 72-77.

Garfield, J. B., & Gal, I. (1999). Assessment and statistics education: Current challenges and directions. *International Statistical Review*, *67*(1), 1-12.

Garfield, J., Hogg, B., Schau, C., & Whittinghill, D. (2002). First courses in statistical science: The status of educational reform efforts. *Journal of Statistics Education*, *10*(2).

Gattuso, L. (2006, July). Statistics and mathematics. Is it possible to create fruitful links. *Proceedings of the Seventh International Conference on Teaching Statistics*. ^ *eSalvador (Bahia) Salvador (Bahia): International Association for Statistical Education and International Statistical Institute.*

Gaughan, J. 2014. "The Flipped Classroom in World History." *History Teacher 47, no. 2: 221-244.*

Gojak L (2012) To flip or not to flip: that is not the question! National Council of Teachers of Mathematics. Retrieved 2016. http://www.nctm.org/about/content.aspx?id=34585

Golafshani, N. (2013) Teachers' beliefs and teaching mathematics with manipulatives. *Canadian Journal for Education*, 36(3), 137-159.

Goodwin, C. (2013). The co-operative, transformative organization of human action and knowledge. *Journal of Pragmatics*, *46*(1), 8-23.

Goodwin, B., & Miller, K. (2013). Evidence on flipped classrooms is still coming in. Educational Leadership, 70(6), 78-80

Goracke, M. A. (2009). The role of manipulatives in the eighth grade mathematics classroom.

Gratz, Z. S., Volpe, G. D., & Kind, B. M. (1993). Attitudes and achievement in introductory statistics classes: Traditional versus computersupported instruction (Report No. JH940166). In *Ellenville, NY: Teaching of Psychology: Ideas and Innovations: Proceedings of the 7th Annual Conference on Undergraduate Teaching of Psychology, March* (pp. 24-26).

Hamden, N., McKnight, P., McKnight, K., & Arfstrom, K. (2013). The flipped learning model: A white paper based on the literature review titled a review of flipped learning. Retrieved October 2, 2013 from http://www.flippedlearning.org/cms/lib07/VA01923112/Centricity/Domain/41/WhitePaper_FlippedLearning.pdf.

Heath, G. D. (2002) Using applets in teaching mathematics. *Mathematics and Computer Education. 36(1), 43-52.*

Herreid, C. F., & Schiller, N. A. (2013). Case studies and the flipped classroom. *Journal of College Science Teaching*, *42*(5), 62-66.

Hilton, S. C., Schau, C., & Olsen, J. A. (2004). Survey of Attitudes Toward Statistics: Factor structure invariance by gender and by administration time. *Structural Equation Modeling*, *11*(1), 92-109.

Hinton-Smith, T. (Ed.). (2012). Widening participation in higher education: casting the net wide? *Palgrave Macmillan.*

Hoaglin, D. C. and Moore, D. S. (eds.) (1992), Perspectives on Contemporary Statistics, MAA *Notes No. 21, Washington, DC: Mathematical Association of America.*

Hogan, K. E., & Pressley, M. E. (1997). Scaffolding Student Learning: Instructional Approaches and Issues. *Brookline Books.*

Hogg, R. V. (1991), "Statistical Education: Improvements are Badly Needed," *The American Statistician*, *45, 342-343*

Hood, M., Creed, P. A., & Neumann, D. L. (2012). Using the expectancy value model of motivation to understand the relationship between student attitudes and achievement in statistics. *Statistics Education Research Journal*, *11*(2).

Huba, M. E., and Freed, J. E. (2000). Using rubrics to provide feedback to students. In: *Learner-Centered Assessment on College Campuses, Boston: Allyn and Bacon, 151–200*

Hubbard, R. (1992). Teaching statistics with MINITAB. *Australian Mathematics Teacher*, *48*(4), 8-10.

Hughes-Hallett, D. (2001). Achieving numeracy: The challenge of implementation. *Mathematics and Democracy: The Case for Quantitative Literacy*, 93-98.

Hulsizer, M. R., & Woolf, L. M. (2009). *A guide to teaching statistics: Innovations and best practices* (Vol. 10). John Wiley & Sons.

Hunt, A. W., Nipper, K. L., & Nash, L. E. (2011). Virtual vs. Concrete Manipulatives in Mathematics Teacher Education: Is One Type More Effective than the Other? *Current Issues in Middle Level Education*, *16*(2), 1-6.

Jacobs, K. L. (2005). Investigation of interactive online visual tools for the learning of mathematics. *International Journal of Mathematical Education in Science and Technology*, *36*(7), 761-768.

Johnson, K. A. (1993). Manipulatives allow everyone to learn mathematics. *Contemporary Education*, *65*(1), 10.

Jones, S. (2003). *U.S. Patent Application No. 10/116,229*.

Keller, B., Wasburn-Moses, J., & Hart, E. (2002). Improving students' spatial visualization skills and teachers' pedagogical content knowledge by using on-line curriculum-embedded applets. *National Council of Teachers of Mathematics Illuminations Project*.

Kelly, C. A. (2006). Using manipulatives in mathematical problem solving: A performance-based analysis. *The Mathematics Enthusiast*, *3*(2), 184-193.

Khan, S. (2011). Let's use video to reinvent education (http://www.ted.com/talks/salman_khan_let_s_use_video_to_reinvent_education.html).

Kirch, C. (2012). Flipping with Kirch. Video retrieved from http://flippingwithkirch.blogspot.ca/

Klahr, D., Triona, L. M., & Williams, C. (2007). Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science teaching*, *44*(1), 183-203.

Klein, D. (2003). A brief history of American K-12 mathematics education in the 20th century. *Mathematical Cognition*, 175-225.

Konold, C. (1995). Issues in assessing conceptual understanding in probability and statistics. *Journal of Statistics Education*, *3*(1).

Konold, C., & Higgins, T. (2003). Reasoning about data. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 193–214). Reston, VA: National Council of Teachers of Mathematics.

Kranzler, Gerald and Janet Moursund. 1999. Statistics for the Terrified. 2d ed. Upper Saddle River, NJ: *Prentice Hall*

Lage, M.J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *Journal of Economic Education*, 31(1), 30-43.

Lindsay, B. G., Kettenring, J., & Siegmund, D. O. (2004). A report on the future of statistics. *Statistical Science*, *19*(3), 387-413.

Lovett, M. C., & Greenhouse, J. B. (2000). Applying cognitive theory to statistics instruction. *The American Statistician*, *54*(3), 196-206.

Love, B., Hodge, A., Grandgenett, N., & Swift, A. W. (2013). Student learning and perceptions in a flipped linear algebra course. *International Journal of Mathematical Education in Science and Technology*, 45(3), 317– 324

Lutzer, D. J., Rodi, S. B., Kirkman, E. E., & Maxwell, J. W. (2007). Statistical abstract of undergraduate programs in the mathematical sciences in the United States: Fall 2005 *CBMS Survey. Washington, DC: American Mathematical Society.*

Marasinghe, M. G., Meeker, W. Q., Cook, D., & Shin, T. S. (1996). Using graphics and simulation to teach statistical concepts. *The American Statistician*, *50*(4), 342-351.

Martin, T. (2009). A theory of physically distributed learning: How external environments and internal states interact in mathematics learning. *Child Development Perspectives*, *3*(3), 140-144.

Marzano, R.J. (1998). A theory-based meta-analysis of research on instruction. *Aurora, CO: Mid-continent Research for Education and Learning.*

Mason, G. S., Shuman, T. R., & Cook, K. E. (2013). Comparing the effectiveness of an inverted classroom to a traditional classroom in an upper-division engineering course. *IEEE Transactions on Education*, *56*(4), 430-435.

McBride, A. B. (1996). Creating a Critical Thinking Learning Environment: Teaching Statistics to Social Science Undergraduates 1. *PS: Political Science & Politics*, *29*(3), 517-521.

Miller, D., Brown, A., & Robinson, L. (2002). Widgets on the web: Using computer-based learning tools. *Teaching Exceptional Children*, *35*(2), 24-28.

Mills, J. D. (2004). Students' attitudes toward statistics: implications for the future. *College Student Journal*, *38*(3).

Minhas, P. S., Ghosh, A., & Swanzy, L. (2012). The effects of passive and active learning on student preference and performance in an undergraduate basic science course. *Anatomical Sciences Education*, *5*(4), 200-207.

Mittag, K. C. (1992, April). Using computers to teach the concepts of the central limit theorem. *Paper presented at the annual American Educational Research Association, San Francisco, CA*

Moore, D. S. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review*, *65*(2), 123-137.

Moore, D. S. (1998). Statistics among the liberal arts. *Journal of the American Statistical Association*, *93*(444), 1253-1259.

Moore, M. (1990). Recent contributions to the theory of distance education. *Open Learning: The Journal of Open, Distance and e-Learning*, *5*(3), 10-15.

Moore, A. J., Gillett, M. R., & Steele, M. D. (2014). Fostering student engagement with the flip. *MatheMatics Teacher, 107*(6), 420-425.

Moyer, P. S. (2001). Are we having fun yet? How teachers use manipulatives to teach mathematics. *Educational Studies in Mathematics, 47*(2), 175-197.

Moyer, P. S., Bolyard, J. J., & Spikell, M. A. (2002). What are virtual manipulatives? *Teaching Children Mathematics*, *8*(6), 372.

Moyer, P. S., & Jones, M. G. (2004). Controlling choice: Teacher, students, and manipulatives in mathematics classrooms. *School Science and Mathematics, 104(1), 16–21.*

Moyer-Packenham, P. S., & Bolyard, J. J. (2016). Revisiting the definition of a virtual manipulative. *International perspectives on teaching and learning mathematics with virtual manipulatives* (pp. 3-23). *Springer, Cham.*

Mvududu, N. (2003). A cross-cultural study of the connection between students' attitudes toward statistics and the use of constructivist strategies in the course. *Journal of Statistics Education*, *11*(3).

National Council of Teachers of Mathematics. (2000). Principles and Standards for School Mathematics. *Reston, VA: Author*

Nielsen, L. 2012. "Five Reasons I'm not Flipping over the Flipped Classroom." *Technology & Learning 32, no. 10: 46.*

Noschese, F. (2011). Khan Academy: My Final Remarks (http://fnoschese.wordpress. com/2011/05/10/khan-academy-my-final-remarks/).

Onwuegbuzie, A. J., DaRos, D., & Ryan, J. (1997). The components of statistics anxiety: A phenomenological study. *Focus on Learning Problems in Mathematics*, 19, 11-35.

Oswald, D. L., & Harvey, R. D. (2003). A Q-methodological study of women's subjective perspectives on mathematics. *Sex Roles*, *49*(3-4), 133-142.

Overmyer, G. R. (2014). The flipped classroom model for college algebra: Effects on student achievement *(Doctoral dissertation, Colorado State University).*

Packard, A. L., Holmes, G. A. y Fortune, J. C. (1993), A comparison of three presentation methods of teaching statistics, *ERIC Document No. 365, 696.*

Perepiczka, M., Chandler, N., & Becerra, M. (2011). Relationship between graduate students' statistics self-efficacy, statistics anxiety, attitude toward statistics, and social support.

Perney, J., & Ravid, R. (1991). The relationship between attitudes towards statistics, math self-efficacy concept, test anxiety and graduate students' achievement in an introductory statistics course. *Unpublished manuscript, National College of Education, Evanston, IL*.

Petocz, A., & Newbery, G. (2010). On conceptual analysis as the primary qualitative approach to statistics education research in psychology. *Statistics Education Research Journal*, *9*(2).

Phyllis (2001) The effects of using computer manipulatives in teaching probability concepts to elementary school students. *EdD Columbia University Teachers College. Supervisor: Brice R. Vogelli*

Pink, D. (2010). Think Tank: Flip-thinking-the new buzz word sweeping the US. The Telegraph

Ramirez, C., Schau, C., & Emmioglu, E. (2012). The importance of attitudes in statistics education. *Statistics Education Research Journal*, *11*(2).

Ramsay, J. O. (2000). Differential equation models for statistical functions. *Canadian Journal of Statistics*, *28*(2), 225-240.

Raphael, D., & Wahlstrom, M. (1989). The influence of instructional aids on mathematics achievement. *Journal for Research in Mathematics Education*, 173-190.

Reed, S. K., & Jazo, L. (2002). Using multiple representations to improve conceptions of average speed. *Journal of Educational Computing Research*, *27*(1), 147-166.

Reimer, K., & Moyer, P. S. (2005). Third-graders learn about fractions using virtual manipulatives: A classroom study. *The Journal of Computers in Mathematics and Science Teaching*, *24*(1), 5.

Reyes-García, V., Kightley, E., Ruiz-Mallén, I., Fuentes-Peláez, N., Demps, K., Huanca, T., & Martínez-Rodríguez, M. R. (2010). Schooling and local environmental knowledge: Do they complement or substitute each other? *International Journal of Educational Development*, *30*(3), 305-313.

Riley, G., Beard, L. A., & Strain, J. (2004). Assistive technology at use in the teacher education programs at Jacksonville State University. *TechTrends*, *48*(6), 47-49.

Roberts, G (2004). Teaching using the Web: conceptions and approaches from a phenomenographic perspective." Advances in research in networked Learning. *Kluwer Academic Publishers, Boston, MA, USA.*

Roshan, S., & Roshan, W. (2012, Aug 24). My view: It's never too late to begin flipping your classroom. [CNN's schools of thought blog]. Retrieved from http://schoolsofthought.blogs.cnn.com/2012/08/24/my-view-its-never-too-late-to-begin-flipping-yourclassroom/

Rowntree, Derek. 2004. Statistics Without Tears: A Primer for Non-Mathematicians. *Upper Saddle River, NJ: Allyn and Bacon.*

Ruddick, K. W. (2012). Improving chemical education from high school to college using a more hands-on approach (Unpublished doctoral dissertation). *University of Memphis, Memphis, TN.*

Rumsey, D. J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education*, *10*(3).

Salkind, Neil J. 2000. Statistics for People Who (Think They) Hate Statistics. *Thousand Oaks*, CA: Sage

Sams, A., & Bergmann, J. (2013). Flip your students' learning. *Educational leadership*, *70*(6), 16-20.

San Pedro, M. O. C. Z., d Baker, R. S., & Rodrigo, M. M. T. (2011, June). Detecting carelessness through contextual estimation of slip probabilities among students using an intelligent tutor for mathematics. In *International Conference on Artificial Intelligence in Education* (pp. *304-311). Springer, Berlin, Heidelberg.*

Sarama, J., & Clements, D. H. (2009). "Concrete" computer manipulatives in mathematics education. *Child Development Perspectives*, *3*(3), 145-150.

Scheaffer, R. L., & Stasny, E. A. (2004). The state of undergraduate education in statistics: A report from the CBMS 2000. *The American Statistician*, *58*(4), 265-271.

Schultz, K. S., & Koshino, H. (1998). Evidence of reliability and validity for Wise's attitude towards statistics scale. *Education and Psychological Measurement, 82, 27–31*.

Sebesta, L. M., & Martin, S. R. M. (2004). Fractions: Building a foundation with concrete Manipulatives. *Illinois Schools Journal*, *83*(2), 3-23.

Shaw, J. M. (2002). Manipulatives enhance the learning of mathematics. *Retrieved, September*, *12*, 2012.

Shaw, S. F. (2009). Transition to postsecondary education. *Focus on Exceptional Children*, *42*(2), 1.

Schau, C. (2003, August). Students' attitudes: The "other" important outcome in statistics education. In *Proceedings of the Joint Statistical Meetings* (pp. 3673-3681).

Schau, C., Stevens, J., Dauphinee, T. L., & Vecchio, A. D. (1995). The development and validation of the survey of antitudes toward statistics. *Educational and Psychological Measurement*, *55*(5), 868-875.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions.

Shaughnessy, J. M. (2007). Research on statistics' reasoning and learning. *Second Handbook of Research on Mathematics Teaching and Learning*, 957-1009.

Snart, J. A. (2010). Hybrid Learning: The Perils and Promise of Blending Online and Face-to-Face Instruction in Higher Education: *The Perils and Promise of Blending Online and Face-to-Face Instruction in Higher Education*. ABC-CLIO.

Sowell, E. J. (1989). Effects of manipulative materials in mathematics instruction. *Journal for Research in Mathematics Education*, 498-505.

Steen, Lynn. 2001. The case for quantitative literacy. In L. A. Steen (Ed), Mathematics and democracy: The case for quantitative literacy. *Princeton, NJ: The National Council on Education and the Disciplines*

Steen, K., Brooks, D., & Lyon, T. (2006). The impact of virtual manipulatives on first grade geometry instruction and learning. *Journal of Computers in Mathematics and Science Teaching*, *25*(4), 373-391.

Strauss. V. 2012. You are so smart … why did you become a teacher? *The Washington Post*, February19. www.washingtonpost.com/blogs/answer-sheet/post/you-are-so-smartwhy-did-you-become-a-teacher/2012/02/19/gIQA2vBNNR_blog.html#pagebreak.

Strayer, J. (2007). The effects of the classroom flip on the learning environment: A comparison of learning activity in a traditional classroom and a flip classroom that used an intelligent tutoring system. *(Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (UMI No. 3279789)*

Strayer, J. (2008). The Effects of the Classroom Flip on the Learning Environment: A Comparison of Learning Activity in a Traditional Classroom and a Flip Classroom that Used an Intelligent Tutoring System. *PhD dissertation. Ohio State University*.

Strayer, J. F. (2012). How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learning Environments Research*, 15(2), 171-193.

Suh, J., Moyer, P. S., & Heo, H. J. (2005). Examining technology uses in the classroom: Developing fraction sense using virtual manipulative concept tutorials. *Journal of Interactive Online Learning*, *3*(4), 1-21.

Sullivan, M. M. (1993). Students Learn Statistics When They Assume a Statistician's Role.

Suydam, M. N., & Higgins, J. L. (1977). Activity-Based Learning in Elementary School *Mathematics: Recommendations from Research*.

Swan, P., & Marshall, L. (2010). Revisiting Mathematics Manipulative Materials. *Australian Primary Mathematics Classroom*, *15*(2), 13-19.

Tarr, J. E., & Lannin, J. K. (2005). How can teachers build notions of conditional probability and independence? *Exploring Probability in School* (pp. 215-238). Springer, Boston, MA.

Tishkovskaya, S., & Lancaster, G. A. (2012). Statistical education in the 21st century: A review of challenges, teaching innovations and strategies for reform. *Journal of Statistics Education*, *20*(2).

Thurstone, L. L. (1970). Attitudes can be measured. In G. F. Summers (Ed.), Attitude measurement (pp. 127–141). *Chicago: Rand McNally*.

Tomlinson, C. A. (2001). How to differentiate instruction in mixed-ability classrooms. ASCD.

Toppo, G. 2011. "Flipped classrooms take advantage of technology. *USA Today*

Toto, R., and H. Nguyen. 2009. "Flipping the Work Design in an Industrial Engineering Course." *Paper presented at 39th ASEE/IEEE Frontiers in education conference, San Antonio, TX, October 18–21.*

Touchton M., (2015), Flipping the classroom and student performance in advanced statistics: evidence from a quasiexperiment, *J. Pol. Sci. Educ., 11, 28–44.*

Triona, L. M., & Klahr, D. (2003). Point and click or grab and heft: Comparing the influence of physical and virtual instructional materials on elementary school students' ability to design experiments. *Cognition and Instruction*, *21*(2), 149-173.

Tucker, B. 2012. The flipped classroom. *Education Next 12(1): 82-83.*

Uttal, D. H., Scudder, K. V., & DeLoache, J. S. (1997). Manipulatives as symbols: A new perspective on the use of concrete objects to teach mathematics. *Journal of Applied Developmental Psychology*, *18*(1), 37-54.

Uttal, D. H., O'Doherty, K., Newland, R., Hand, L. L., & DeLoache, J. (2009). Dual representation and the linking of concrete and symbolic representations. *Child Development Perspectives*, *3*(3), 156-159.

Ültanir, E. (2012). An epistemologic glance at the constructivist approach: Constructivist learning in Dewey, Piaget, and Montessori.

Velleman, P. F. (2008). Truth, damn truth, and statistics. *Journal of Statistics Education*, *16*(2).

Velleman, P. F., & Moore, D. S. (1996). Multimedia for teaching statistics: Promises and pitfalls. *The American Statistician*, *50*(3), 217-225.

Vinson, B. M. (2001). A comparison of preservice teachers' mathematics anxiety before and after a methods class emphasizing manipulatives. *Early Childhood Education Journal, 29*(2), 89-94.

Walkington, C., Cooper, J., & Howell, E. (2013). The effects of visual representations and interest-based personalization on solving percent problems. In *Proceedings of the 35th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 533-536). *Chicago, IL: University of Illinois at Chicago.*

Walvoord B.E & Anderson, V,J Effective Grading: A Tool for Learning and Assessment, *Jossey-Bass Publishers, San Francisco, 1998.*

Wang, S., Jiao, H., Young, M. J., Brooks, T., & Olson, J. (2007). A meta-analysis of testing mode effects in grade K-12 mathematics tests. *Educational and Psychological Measurement*, *67*(2), 219-238.

Watson, J. M. (2013). *Statistical literacy at school: Growth and Goals*. Routledge.

Watson, L. (2006). The role of spirituality and religion in the experiences of African American male college students. *M. J. Cuyjet (Ed.), African American men in college (pp. 112-127). San Francisco, CA: Jossey-Bass.*

Williams, A. S. (2013). Worry, Intolerance of Uncertainty, and Statistics Anxiety. *Statistics Education Research Journal*, *12*(1).

Wilson, S. (2013). Investigating rural pre-service teachers' mathematics anxiety using the Revised Mathematics Anxiety Scale (RMARS). *Australian and International Journal of Rural Education*, *23*(3), 1.

Young, S. (2006). Student views of effective online teaching in higher education. *The American Journal of Distance Education*, *20*(2), 65-77.

Zappe, S., Leicht, R., Messner, J., Litzinger, T., & Lee, H. W. (2009). " Flipping" the classroom to explore active learning in a large undergraduate course. In *ASEE Annual Conference and Exposition, Conference Proceedings*.

Zeidner, M., & Shani-Zinovich, I. (2011). Do academically gifted and non-gifted students differ on the Big-Five and adaptive status? Some recent data and conclusions. *Personality and Individual Differences*, *51*(5), 566-570.

Zinn, T. E., & Smiley, W. F. (2003) Activities for Engagement in an Industrial/Organizational Psychology Course. *Promoting Student Engagement*, 54.

ANUSHKA KARKELANOVA

VITA

Professional Experience

01/2018-Present University of Kentucky
Lecturer, Department of Statistics

08/2012 –12/2017 University of Kentucky
Full Time Instructor, Department of Statistics

07/2008 to 06/2012 Woodford County High School
Math Teacher 9^{th}-12^{th} Grade

07/2007 to 06/2008 Tates Creek High School
Math Teacher 9^{th} Grade

08/2006 to 06/2007 Lexington Traditional Magnet School
Math Teacher 7^{th} Grade

01/2006 to 05/2007 UK Department of Special Education
Research Assistant
- Math Pursuits Grant

08/2004 to 08/2006 UK Department of Statistics and LLC Lexington, KY
Part Time Instructor

08/2001 to 08/2004 UK Department of Statistics and LLC Lexington, KY
Teaching Assistant

Education and Certifications

- 1998 Technical University Sofia branch Plovdiv-Master in Electronic Engineering Computer Science
- 2004 University of Kentucky-Master in Statistics
- 2007-2017 Certificate for College Faculty-Rank I, Mathematics, Grades 8-12
- Certificate for Computer Operator and Programmer Qualification (Pascal, Basic, Fortran)

Accomplishments

- July 2019- USA coordinator for International Statistical Literacy Project
- Spring 2017-Faculty Fellow Leadership Team Member –Cohort 6, University of Kentucky honor.
- January 2015-CESL Star, University of Kentucky honor.
- SAGE Publisher-Statistics Education Book Reviewer 2014
- July 21-25th 2008 -Middle Grades Textbook Reviewer Evaluation process.