Self-efficacy and Collaborative Learning: An Intervention Study

Jane Robertson

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Education.

Chapel Hill
2012

Approved by:
Jeffrey A. Greene, Ph.D., Chair
Susan Friel, Ph.D.
Jill Hamm, Ph.D.
Jill Thomley, Ph.D.
William B. Ware, Ph.D.
ABSTRACT

JANE ROBERTSON: Self-efficacy and Collaborative Learning: An Intervention Study
(Under the direction of Jeffrey A. Greene)

Findings from empirical research suggest that both self-efficacy beliefs and collaborative learning may have an influence upon student academic performance. However, the phenomena of self-efficacy beliefs, collaborative learning, and academic achievement have not been studied in concert with one another. Using quantitative research methods, I sought answers to the following research questions:

1. Does working in small, collaborative, student led, peer groups in the classroom increase the topic specific self-efficacy beliefs of undergraduate level statistics students?

2. Do topic specific student self-efficacy beliefs mediate the relationship between collaborative learning and the academic achievement of undergraduate level statistics students studying statistical sampling?

I administered pre and post surveys and assessments to establish answers to the two research questions above. I also employed qualitative research methodology involving iterative cycles of observation and reflection, as well as retrospective analysis, to elucidate areas of the collaborative intervention task design that required modification for improvement both during and after the study. Using these qualitative methods, I was able to answer the two final research questions:
3. How should the collaborative tasks intended to help undergraduate students learn topics related to statistical sampling be redesigned and implemented such that they foster the topic specific self-efficacy beliefs and, thereby, promote higher academic achievement of undergraduate students?

4. How should the overall research design be revised for use in future experiments of the same type implemented in similar contexts?

In the present study, I was not able to establish support for the hypothesis that working in small, collaborative groups in the classroom positively affected students’ topic specific self-efficacy beliefs. Nor was I able to show that self-efficacy beliefs mediated the relationship between collaborative learning and academic achievement in statistical sampling for this group of students. However, through the use of qualitative research methods, I was able to uncover potential explanations for why I was unable to establish support for these hypotheses. I was able to determine how to modify the collaborative learning tasks and the overall research design in ways that strengthened these components for use in future research studies. These changes included developing collaborative learning interventions with consideration of the architectural layout of the classroom, effectively aligning the use of various instructional methods with students’ needs, and allowing ample time for students to become comfortable working in a collaborative learning environment. I believe these improvements will positively impact the self-efficacy beliefs and academic achievement of students engaged in future collaborative learning interventions.
ACKNOWLEDGEMENTS

This work would not have been possible without the support of a very special group of people who loved and supported me along the way. First, I would like to thank my committee for their guidance throughout this process. I would especially like to thank my mentor and advisor, Jeff Greene, for always being there when I needed him and saying just the right thing. There is no one with whom I would rather have navigated this journey. At times the lessons were painful, but we did in fact learn together. ‘Thank you’ seems so small.

To my cohort, thank you for teaching me how to create and foster a loving, supportive community of people who began strangers. We have forged a bond that we will carry with us. I look forward to continuing to share the experience of becoming who we are supposed to be. Debbie and Jenn, I truly believe I would not have completed this without you, nor would I have wanted to.

I would like to thank my family, those to which I was born and those whom I have chosen. Mom, you taught me the art of moving through life as a strong, independent woman. For that, I am forever grateful. Susan, I’m not sure others understand the incredibly strong bond we share. Thank you for teaching me to read and so much more. To April, Jonathan, Lorna, and Mary, thank you for choosing me and allowing me to be part of your lives. There are no words to express the immense love and gratitude I feel for each one of you. Finally, to Dad - I lost your physical presence
during this pilgrimage, yet you were with me all along. Thank you for raising two strong daughters who, because of you, believe they can do anything.
# TABLE OF CONTENTS

LIST OF TABLES ...........................................................................................................x

LIST OF FIGURES .........................................................................................................xi

Chapter

I. INTRODUCTION ...........................................................................................................1
   Self-efficacy Beliefs ......................................................................................................3
   College Statistics ........................................................................................................5
   Collaborative Learning ...............................................................................................8
   The Effects of Collaborative Learning on Self-efficacy Beliefs .........................12
   Purpose of Study .........................................................................................................14
   Research Questions .....................................................................................................15
   Contributions of Present Study ..................................................................................16

II. LITERATURE REVIEW .............................................................................................18
   Self-efficacy Beliefs ....................................................................................................19
   Statistics Education ...................................................................................................32
   Collaborative Learning ..............................................................................................41
   Intervention Research ...............................................................................................56
   Conclusion ..................................................................................................................61

III. METHODS ..............................................................................................................64
   Participants ................................................................................................................65
APPENDIX E........................................................................................................226
APPENDIX F........................................................................................................229
APPENDIX G........................................................................................................238
APPENDIX H........................................................................................................248
APPENDIX I........................................................................................................253
APPENDIX J........................................................................................................257
REFERENCES......................................................................................................261
LIST OF TABLES

Table

1. Student Demographics.................................................................67
2. Student Attendance........................................................................68
3. A Priori Power Analysis..................................................................95
4. Descriptive Statistics.................................................................100
5. Direct and Indirect Effects – Original Path Model......................102
6. Direct and Indirect Effects – Final Path Model............................106
7. Threats to Internal Validity..........................................................112
LIST OF FIGURES

Figure

1. Hypothesized path model.................................................................94
2. Original path model with estimates..................................................101
3. Interaction path model with estimates..............................................104
4. Final path model with estimates.....................................................105
5. Graphical representation of interaction...........................................109
CHAPTER 1

INTRODUCTION

According to the National Center for Education Statistics, only 57% of undergraduate students who began a degree program in 2001 at a 4-year college or university received a bachelor’s degree from that institution within six years (U. S. Department of Education, 2010). Developing innovative ways to foster academic achievement, thus promoting college degree completion, has been an ongoing challenge for researchers, practitioners, and administrators alike (U. S. Department of Education, 1983). Based on research in the field of self-efficacy, or, “beliefs that people hold about their capability to succeed in their endeavors” (Pajares, 2006, p. 339), it has been shown that students with higher self-efficacy beliefs for a particular subject, and even for a specific task, tend to achieve at higher levels within that domain or on that task than students with lower levels of self-efficacy beliefs (Bandura, 1997; Pajares, 1996; Pajares & Urdan, 2006). Self-efficacy scholars have suggested that educators must understand how to bolster students’ self-efficacy beliefs as a means of creating academic success (Chemers, Li-tze, & Garcia, 2001; Pajares, 1996). While many studies have focused on various methods of promoting the self-efficacy beliefs of middle and high school students, the empirical research regarding the techniques instructors can employ to help college students increase their self-efficacy beliefs is inadequate. Frank Pajares (1996), a scholar who primarily studied the self-efficacy beliefs of adolescents in schools, suggested that further research is warranted to understand how students at all academic levels, including college
students, are affected by the various phenomena that are known to have an effect upon students’ sense of self-efficacy. One of the primary phenomena that influences student self-efficacy is peer modeling, or the act of students observing their peers successfully complete a task through effortful perseverance (Bandura, 1995). Based upon a thorough review of the research literature, it seems that self-efficacy scholars have not heeded Pajares’ (1996) advice to study peer modeling among college students. Further empirical research is required to determine how to positively influence college students’ self-efficacy beliefs and, thus, promote academic achievement and degree completion.

I believe that collaborative learning is one method of instruction that can help college students increase their domain and topic specific self-efficacy beliefs, thereby promoting higher academic achievement. During collaborative learning exercises, students are able to observe their peers working to successfully complete academic tasks (Johnson & Johnson, 1989). The peer modeling aspect of collaborative learning can, in turn, increase students’ self-efficacy beliefs, leading to greater academic success (Bandura, 1995). Therefore, I believe that self-efficacy beliefs mediate the relationship between collaborative learning and academic achievement in statistical sampling. Further research is needed to establish these causal relationships among collaborative learning, students’ self-efficacy beliefs, and academic achievement.

In this study, I implemented a collaborative learning intervention using both quantitative and qualitative research methods. The effects of the collaborative learning intervention were evaluated using quantitative methods including pre and post intervention surveys and assessments. In an attempt to increase the efficacy of the collaborative learning treatment and to gather practical data that may be beneficial to
educators who wish to implement similar interventions in their classrooms, the qualitative methodology of the present study involved iterative cycles of classroom observation and reflection, as well as retrospective analysis, to inform both the intervention as it was implemented as well as future instances of collaborative learning interventions. The qualitative research methodology employed in this study was informed by the ideas of design research (Gravemeijer & van Eerde, 2009). I hypothesized that the results of the present study would provide educators with a springboard from which to create their own unique collaborative classroom environment in which students are provided opportunities to learn in an atmosphere that can help foster increased self-efficacy beliefs and potentially advance their academic success.

**Self-Efficacy Beliefs**

Albert Bandura is regarded as the father of self-efficacy theory (Pajares & Urdan, 2006). Bandura (1997) defined self-efficacy as, “people’s beliefs in their capabilities to produce desired effects by their actions” (p. vii). Bandura (1977) began his work on what would become known as self-efficacy theory in the late 1970s and continued to study this topic throughout his career. His ideas regarding self-efficacy beliefs began with his social cognitive theory (Bandura, 1986). A key point of social cognitive theory is that self-reflective thought allows individuals to assess their knowledge, experiences, and thoughts. Bandura (1977) maintained that, through self-reflection, people appropriately alter their thoughts and behaviors. He believed that people use what he called *reciprocal determinism* to evaluate their thoughts and make the adjustments necessary to integrate information into their collections of knowledge and experiences. The idea of reciprocal determinism is that personal aspects (i.e., cognitive, affective, and biological
experiences), people’s behavior, and influences from the environment interact with, and are influenced by one another. According to Bandura, these interactions prompt people to refine their thinking and actions. Self-efficacy beliefs are adjusted during the thought evaluation and refinement process.

Self-efficacy research based on Bandura’s (1977) theory has been conducted across many academic domains and with various age groups, including adolescents in general academic settings (Pajares & Urdan, 2006), middle-school students studying mathematics (Pajares & Miller, 1997), college students enrolled in statistics courses (Finney & Schraw, 2003), and first year college students in relation to academic performance and emotional adjustment (Chemers et al., 2001). Because Bandura’s (1977) self-efficacy theory is the principal theory accepted in the field, it also served as a foundation for the present intervention study.

There has been much debate in the field of self-efficacy regarding the level of specificity of student self-efficacy beliefs (Pajares & Urdan, 2006). That is, scholars have questioned whether students have a general sense of academic self-efficacy (Chemers et al., 2001) or if self-efficacy beliefs are specific to various domains and, possibly, to a set of tasks or topics within each domain (Pajares, 1996). I believe that students hold overarching beliefs about their success as a student in general as well as specific self-efficacy beliefs that differ by domain and by task or topic within each domain (Bandura, 1997; Pajares, 2006). The present intervention study focused on students’ self-efficacy beliefs in relation to the various topics associated with statistical sampling within the domain of introductory undergraduate statistics.
Self-efficacy scholars have conducted numerous research studies to investigate the relationship between student self-efficacy beliefs and academic achievement, in the hopes of providing evidence in support of the idea that increased self-efficacy beliefs lead to higher academic performance among students at all levels of education. The research has overwhelmingly established that increased self-efficacy beliefs, whether generally academic or domain, topic, or task specific, are in fact positively related to academic performance (see Chemers et al., 2001; Finney & Schraw, 2003; Pajares & Miller, 1997). However, specific intervention research is needed, especially in the college classroom, to help educators understand how to bolster students’ academic, domain specific, and topic or task specific self-efficacy beliefs. I believe that collaborative learning is an instructional tool that college practitioners can utilize in their classrooms to help students make positive changes to their self-efficacy beliefs that, in turn, can increase academic achievement (Wang & Lin, 2007).

**College Statistics**

The transition from high school to college can be a daunting time for many students (Zarrett & Eccles, 2006). Beyond the social challenges students face when they begin college, many also find the academic adjustment to be difficult (Chemers et al., 2001). Chemers and colleagues found that academic self-efficacy and optimism were significantly positively related to first-year students’ adjustment to college and to their academic performance. That is, increased self-efficacy beliefs may help college students navigate the transition from high school, both socially and academically (Chemers et al., 2001), which can increase the likelihood that they will remain in college and complete an undergraduate degree (DeWitz, Woolsey, & Walsh, 2009). Within the domain of
statistics, further experimental research is necessary to establish the effect of students’ self-efficacy beliefs upon adjustment to college and academic performance in the statistics classroom, contributing to overall collegiate achievement and degree completion.

The concepts taught in introductory and advanced level statistics courses are often challenging topics for students (Garfield & Ahlgren, 1988). When college students enter the statistics classroom, both at the undergraduate and graduate levels, for many it is the first time they have heard the terms *standard deviation, hypothesis test,* or *confidence interval* (Ben-Zvi & Garfield, 2005). Students are often anxious about their academic performance in this particular domain. The Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report (2005), a report compiled by a group of leaders in the field of statistics education, included six recommendations for teaching introductory statistics in the college classroom. The first of these recommendations was that instructors should, “emphasize statistical literacy and develop statistical thinking” (GAISE, 2005, p. 1). Statistical literacy has been defined as, “basic and important skills that may be used in understanding statistical information or research results” (Ben-Zvi & Garfield, 2005, p. 7). These basic and important skills include data organization skills, the ability to create accurate tables and graphs, and a basic understanding of statistical ideas and language. Statistical thinking, or reasoning, is a bit more complex than statistical literacy. According to Ben-Zvi and Garfield (2005), statistical thinking encompasses the ways in which students understand statistical data and information. Statistical thinking skills include the ability to make reasonable interpretations from data,
to make cognitive connections between different statistical ideas, and to effectively explain the understanding of various statistical concepts.

The topics associated with statistical sampling can be overly abstract and complex for introductory undergraduate statistics students (Nguyen, 2005; Yilmaz, 1996). Gaining a full understanding of these topics requires that students possess analytical skills that can be difficult to teach (Yilmaz, 1996). Because introductory statistics students typically have varying mathematical and statistical backgrounds, instructors often struggle to teach the material associated with statistical sampling (Fecso et al., 1996; Yilmaz, 1996). For many undergraduate students, a course in statistical sampling or an introductory course in which statistical sampling is a topic covered over a period of a few weeks may be their only exposure to the subject. This ‘once-in-a-lifetime’ experience for students who will not become experts in a field requiring deep understanding of the topics associated with statistical sampling creates a situation in which instructors often struggle to determine how much academic material to cover (Fecso et al., 1996; Yilmaz, 1996).

There is a preponderance of empirical evidence suggesting that collaborative group work is one of the best teaching and learning methods that instructors can employ to help students become more statistically literate and enhance statistical reasoning skills, thus gaining a better understanding of the subject matter and developing higher-order thinking in this domain (e.g., Garfield, 1995; Garfield & Ahlgren, 1988; Jordan, 2007; Lovett, 2001; Roseth, Garfield, & Ben-Zvi, 2008; Zieffler et al., 2008). However, the mechanism by which collaborative learning increases academic achievement is not yet fully understood. In the present study, I investigated this relationship and attempted to
establish student self-efficacy beliefs as a mechanism that mediates the relationship between collaborative learning and academic achievement for college statistics students studying sampling.

**Collaborative Learning**

Collaborative learning is often confused with other types of classroom learning, such as cooperative learning or simply students working together in small groups. Abrami and colleagues (1995) defined cooperative learning as, “an instructional strategy in which students work together in groups that are carefully designed to promote positive interdependence” (p. 1). While collaboration assumes cooperation, the important distinction between the two types of learning is that collaboration requires students to work together, negotiate, argue their points, use conflict resolution skills, and produce a single final product that reflects the ideas and opinions of the group as a whole. Students working together in a cooperative learning environment may each produce their own, individual final product, or may produce a final product that does not reflect the ideas or opinions of the entire group. Palinscar and Herrenkohl (1999) stated that, “the essence of collaboration is convergence or the construction of shared meanings for conversations, concepts, and experiences” (p. 152). A collaborative group’s end product can assume a variety of formats and can be either tangible or intangible. For example, a collaborative group may work toward composing a thoughtful body of knowledge, constructing a glossary of terms, building a portfolio, or developing an authentic performance of some nature. Regardless of which form the final product takes, the key element of collaborative learning is that the end result is a piece of work that is produced with input from all members of the group and is one common, agreed upon creation.
The Benefits of Collaborative Learning

Cobb (1992) said it well when he wrote, “shorn of all subtlety and led naked out of the protective fold of educational research literature, there comes a sheepish little fact: lectures don’t work nearly as well as many of us would like to think” (p. 9). While some instructors continue to use traditional methods of instruction including lecture, rote memorization, and ‘drill and kill’, the use of small, student led, collaborative groups as a means of promoting student learning, social skills, and academic achievement has become increasingly popular in the classroom (Webb, 2009). One of the key benefits of collaborative learning is the socially interdependent environment created by the students (Johnson & Johnson, 1989).

Many scholars who have studied collaborative learning have employed social interdependence theory as the theoretical foundation for their research (Johnson & Johnson, 2002). Johnson and Johnson (1989) defined social interdependence as a learning environment where the behavior, achievements, and academic success of each individual member are affected by the actions of all other group members. They noted that the term *interdependence* implies a bidirectional relationship between and among members of a collaborative learning group, as opposed to a one-way dependent relationship or completely independent working environment. Johnson and Johnson explained the bidirectional relationship between social interdependence and group interactions in the following manner: “cooperation promotes trust, trust promotes greater cooperation, which promotes greater trust, and so forth” (p. 9).

According to Johnson and Johnson (1989), the socially interdependent environment created within collaborative learning groups allows peer modeling to occur.
The authors suggested that, when students work together in small, collaborative groups and have the opportunity to observe their peers working to achieve academically, they often model their successful peers. According to Johnson and Johnson, this advantage of collaborative learning is not typically present when a non-peer educator is leading instruction. As mentioned previously, peer modeling is one of the primary sources through which students increase their self-efficacy beliefs (Bandura, 1995). Thus, it follows that collaborative learning can positively affect students’ self-efficacy beliefs, leading to higher academic achievement among students at all levels of learning, including high achievers and students who typically perform at moderate and lower academic levels as well (Johnson & Johnson, 1989).

**Considerations When Developing a Collaborative Learning Intervention**

There are several aspects that should be considered when developing a collaborative learning intervention. Cohen (1994b), through her work with middle school students, presented several suggestions that she believed helps create productive, collaborative student groups. In the present study, I followed these recommendations and took into consideration the following key principles:

1. Instructors must prepare students to work collaboratively. Cohen noted that not all students know how to interact with their peers in a collaborative learning group and that it is up to the instructor to ensure that students are sufficiently prepared to engage in this type of learning environment. She suggested that the instructor spend time teaching students the social skills and norms necessary to participate in the type of collaborative groupwork to be implemented. Cohen recommended that
these norms and skills be taught through games and exercises she called ‘skillbuilders’.

2. Students should be grouped heterogeneously so that group members can use their peers as resources. Cohen noted that the best groups are those comprised of low and middle level achievers or middle and high level achievers. She suggested this arrangement to avoid frustration on the part of students who are at vastly different academic levels (i.e., low and high level achievers). Cohen also cautioned instructors to avoid grouping students homogeneously, as this type of grouping is unlikely to provide sufficient peer resources for low achieving students.

3. Students should be provided with written instructions for each task. Cohen believed that, after the instructor has presented verbal instructions for the collaborative task, written instructions provide students with a guide to follow during collaborative learning sessions. She noted that written instructions reduce the need for outside assistance from the instructor.

4. Students should be arranged in groups of four to five. According to Cohen, in a group of more than five students, it is likely that one or more students will be left out of the conversation and activity. Likewise, in a group of three, two students may form a bond and leave the third student to work on their own. And a pair of students does not provide enough peer feedback, varying perspectives, and resources.
5. Students should be seated at a table such that everyone can see and hear their peers. Cohen suggested that the ideal arrangement is to seat students in a circle. She also noted that student groups should be seated far enough away from one another so that groups will not disturb each other.

6. Each collaborative task or class period should end with a ‘wrap-up’ period. Cohen believed that this final debriefing period is essential to ensuring that students understand the material, are allowed to share their ideas, and have the opportunity to clear up any misconceptions.

7. Each student in the group should be assigned a clearly defined role. For example, Cohen suggested that each group should have a facilitator, materials manager, reporter, checker, etc. She attested that assigning roles to students will ensure that they engage in collaboration and work through the tasks more efficiently than if students are allowed to interact without these role assignments.

The Effects of Collaborative Learning on Self-efficacy Beliefs

Both increased self-efficacy beliefs and working together with peers in small collaborative classroom groups can lead to higher academic achievement (Bandura, 1997; Cohen, 1994a; Johnson & Johnson, 1989). The phenomena of collaborative learning and self-efficacy have been studied jointly in the past, although the volume of research considering the two together is meager at best. One line of research in this area has focused on how students’ self-efficacy beliefs affect their performance in small group settings (Ruys, Van Keer, & Aelterman, 2010; Sins, van Joolingen, Savelsbergh, & van Hout-Wolters, 2008; Wang & Lin, 2007). In a study conducted by Wang and Lin (2007),
college students enrolled in an introductory educational psychology course were
separated into three categories based upon self-reported self-efficacy beliefs about the
course. The three groups were:

1. Students with high self-efficacy beliefs
2. Students with low self-efficacy beliefs
3. A mixed group including students whose self-efficacy beliefs fell between low
   and high, as well as randomly selected students from both the low and high
   groups

Wang and Lin found that the group comprised of students with high self-efficacy beliefs
for this educational psychology course had higher collective efficacy (i.e., individual
beliefs about the achievement ability of the group) and employed higher-order thinking
more often than did the other two groups.

There is little to no research, however, centered around how working in small
collaborative groups in the classroom affects student self-efficacy beliefs regarding a
specific academic topic. In their concluding remarks, Wang and Lin (2007) suggested
that practitioners should place at least one student with high topic specific self-efficacy
beliefs within each collaborative learning group. The authors linked this suggestion back
to the peer modeling aspect of Bandura’s (1986) self-efficacy theory. Bandura suggested
that observing a similar peer (i.e., a model) who exhibits high academic achievement and
strong self-efficacy beliefs can help students with lower self-efficacy beliefs increase
their self-efficacy and, in turn, achieve greater academic success. Wang and Lin (2007)
echoed Bandura by stating that, “students with high efficacy beliefs not only have
modeling effects on other group members, but are also more likely to transmit their efficacy beliefs through interactions with others” (p. 2265).

Investigating the effects of collaborative learning on self-efficacy beliefs would greatly enhance the overall body of literature and advance the understanding of both fields of study. The favorable relationship shared between collaborative learning and students’ academic achievement may be fostered by the positive effects peer modeling has upon students’ self-efficacy beliefs. Further study of these relationships will provide an understanding of the role self-efficacy, and specifically peer modeling, plays in the relationship between collaborative learning and academic success.

**Purpose of Study**

In the present study, I aimed to determine the mechanism by which collaborative learning positively affects college students’ academic performance in the domain of statistics, specifically the task of learning about statistical sampling. I attempted to establish students’ topic specific self-efficacy beliefs as the mediating factor between the collaborative learning intervention and academic achievement, which in this study was the learning of statistical sampling. In order to establish the causality of the relationships among these phenomena, I conducted a quasi-experimental intervention study. This type of research study allows for the assertion of such causal claims, whereas a non-experimental study does not (Campbell & Stanley, 1963).

To ensure that the intervention was administered as intended, as well as to offer suggestions to researchers and practitioners who may want to conduct future studies of this type, I employed qualitative research methodology to study the treatment group of students. By observing students as they worked together in a collaborative learning
environment, and using these daily observations along with my personal reflections upon the experiences of each class period, I was able to iteratively alter the collaborative learning intervention tasks in a real-time manner. The qualitative data gathered during the classroom intervention brought to light various modifications for improvement of the daily intervention tasks over the course of the study and helped increase the efficacy of the collaborative learning treatment.

While quasi-experimental intervention studies are often quantitative in nature, it has been suggested that using a mixed methods approach can provide richer, more meaningful data (Pressley, Graham, & Harris, 2006). Collecting both qualitative and quantitative data in the present study provided a deep understanding of what the students in the treatment class learned and how they received and experienced the collaborative learning intervention. The combination of qualitative and quantitative data allowed me to make recommendations that serve to improve future research studies and instructional interventions of this type.

**Research Questions**

In this study, I sought to provide an understanding of how collaborative learning affects the academic achievement of students studying statistical sampling by answering the following research questions:

1. Does working in small, collaborative, student led, peer groups in the classroom increase the topic specific self-efficacy beliefs of undergraduate level statistics students?
2. Do topic specific student self-efficacy beliefs mediate the relationship between collaborative learning and the academic achievement of undergraduate level statistics students studying statistical sampling?

3. How should the collaborative tasks intended to help undergraduate students learn topics related to statistical sampling be redesigned and implemented such that they foster the topic specific self-efficacy beliefs and, thereby, promote higher academic achievement of undergraduate students?

4. How should the overall research design be revised for use in future experiments of the same type implemented in similar contexts?

**Contributions of the Present Study**

As a means of helping college statistics students gain sound knowledge of statistical topics and achieve academically, researchers and practitioners need to fully understand how to help students increase their topic specific self-efficacy beliefs. I believe that the peer modeling aspect of collaborative learning may be one way to help students bolster these beliefs. Further, until recently, much of the research concerning collaborative learning has focused on K-12 students, primarily those students in elementary and middle schools (Johnson & Johnson, 2009). Due to students’ perceived unwillingness to participate, the architectural setup of a typical college classroom, and the competitive nature of college itself (Cohen, 1994a), the college classroom has not typically been viewed as an arena in which to use collaborative learning as an instructional tool. The research literature is also lacking in studies focusing on the self-efficacy beliefs of college statistics students. Because the present intervention study use both quantitative and qualitative methods to examine the self-efficacy beliefs of college
level statistics students working collaboratively, it not only fills several gaps in the current literature, but it also provides new directions for future research and practice.

This intervention study provides an understanding of how to design, implement, and redesign collaborative tasks for undergraduate statistics students. The qualitative research methods used in this study provide insight into what types of tasks may or may not be appropriate for college students in statistics classes. The qualitative data gathered also helped to determine how the overall research study was best revised and strengthened. Because the results of the present study are not intended to be applied to various contexts, but instead are specific to the context in which the tasks were implemented, this study does not provide general suggestions for the use of collaborative tasks in any undergraduate statistics course. However, the results of this study provide instructors and researchers a foundation from which to develop their own research study or intervention, specific to their particular context.
CHAPTER 2
LITERATURE REVIEW

There is a wealth of literature regarding student self-efficacy beliefs, collaborative learning, and the difficulties faced by students in the statistics classroom. The literature examining these phenomena in concert, however, is severely lacking. In the present intervention study, I investigated the relationship between collaborative learning and student self-efficacy beliefs in the domain of college introductory undergraduate statistics. Developing a clear understanding of introductory statistics, and thus attaining achievement within a statistics course, has been a challenge faced by many college statistics students (Garfield, 1995). I hypothesized that the peer modeling aspect of collaborative learning can help increase topic specific self-efficacy beliefs, and therefore academic achievement, of college statistics students learning about statistical sampling. The topics associated with statistical sampling, such as data collection, research design, sampling methods, and survey development can be overly complex and abstract for undergraduate students taking an introductory course (Fecso et al., 1996; Gelman & Nolan, 2002; Nguyen, 2005; Utts & Heckard, 2006; Yilmaz, 1996). In order to fully understand the effects of the treatment of this intervention study, I chose to use qualitative research methodology to study the relationships among collaborative classroom learning, students’ self-efficacy beliefs, and academic achievement of the students in the treatment class learning about statistical sampling. In this section, I
present a review of the literature regarding self-efficacy beliefs, statistics education with specific focus on the challenges of learning statistical sampling, collaborative learning, and intervention research.

**Self-efficacy Beliefs**

**The Work of Albert Bandura**

Albert Bandura (1977) was the first academic scholar to present the idea that positive self-efficacy beliefs can promote favorable outcomes. He began his work with self-efficacy theory in the field of psychology, eventually applying his theory to other domains, including education (Bandura, 1986). Bandura (1995) believed there are four main factors that influence self-efficacy beliefs: mastery experiences, vicarious experiences, social persuasion, and affective states. Mastery experiences are based on students’ past experiences with a subject or specific topic or task. If students have had success in the past with a particular task, they are more likely to have higher self-efficacy beliefs about completing a similar task at present or in the future. On the other hand, if students have experienced failure with a task similar to the one currently presented, they will likely have lower self-efficacy beliefs regarding their ability to successfully complete the task.

Bandura (1995) believed that vicarious experiences can also help increase student self-efficacy beliefs. He stated that, “seeing people similar to themselves succeed by perseverant effort raises observers’ beliefs that they, too, possess the capabilities to master comparable activities” (Bandura, 1995, p. 3). While Bandura believed that mastery experience has the most influence on self-efficacy beliefs, the process of observing a model (i.e., a similar peer) work diligently and reach high academic
achievement has been shown to be very influential as well (Cohen, 1994a; Johnson & Johnson, 1989; Pajares 1996). For example, Schunk and Hanson (1985) found that observing high achieving peer models had a statistically significant positive effect on the topic specific self-efficacy beliefs of elementary school children who were learning subtraction skills.

The third factor that Bandura (1995) said can influence self-efficacy beliefs is social persuasion. Using persuasive comments to increase the self-efficacy beliefs of someone who truly does possess academic potential for success, but may not realize it, can be an effective tool. However, false comments stated as a means to increase the self-efficacy beliefs of someone who does not actually have the ability to achieve can have an adverse effect. Bandura believed that people can see through the disingenuous nature of such comments.

Finally, affective states are the fourth factor that can influence self-efficacy beliefs (Bandura, 1995). Affective states include stress, tension, and positive or negative moods and physical states such as fatigue, aches, and pains. Improving the perception of physical and affective states can help increase self-efficacy beliefs. Simply stated, being in a good mood and physically feeling well will likely produce higher self-efficacy beliefs than feeling depressed or ill (Dierdorff, Surface, & Brown, 2010; Kwan & Bryan, 2010).

Usher and Pajares (2008) conducted a review of the self-efficacy literature, focusing on research studies that investigated one or more of the four sources of self-efficacy beliefs put forth by Albert Bandura (1995). Usher and Pajares (2008) limited their literature search to studies that were conducted in schools. They identified key
findings, exposed methodological problems in several of the studies and gaps in the literature, and made suggestions for future research in this area.

Through their literature review, Usher and Pajares (2008) found that mastery experience regularly predicted self-efficacy beliefs. This held true across various domains and for all grade levels. The three remaining sources of self-efficacy presented by Bandura (1995) did not always predict self-efficacy beliefs in a consistent manner. Usher and Pajares (2008) believed this was due to problems with research methodology, the use of measurement instruments that produced unreliable data, and contextual issues they identified in various studies. For example, the review of literature revealed that when researchers summed the values of individual self-efficacy items to create an overall score, the aggregated scores often obfuscated the effects of each individual source of self-efficacy. Usher and Pajares also found that the results of several studies included in their review indicated multicollinearity between the factors that Bandura (1995) believed influence self-efficacy. Regarding contextual issues, Usher and Pajares (2008) found that, “the predictive value of the sources depend on the domain in which the constructs are assessed, and both their magnitude and their relationship with self-efficacy are influenced by students’ group memberships or academic ability indexes” (p. 781). They suggested that further research is necessary regarding vicarious experience, persuasive comments, and affective states, to truly determine if these are means by which self-efficacy beliefs can be positively affected.

Albert Bandura spent his career furthering the understanding of self-efficacy beliefs (Pajares, 1996). His work has provided a foundation upon which contemporary scholars have based their empirical research regarding the phenomenon of student self-
efficacy beliefs. While Bandura’s (1977) theory of self-efficacy is widely accepted throughout the field, there remains debate over the level of specificity of student self-efficacy beliefs.

**The Specificity Controversy**

Some would argue that self-efficacy beliefs are domain specific and possibly even task or topic specific (Finney & Schraw, 2003; Pajares, 1996; Pajares & Miller, 1997). Others believe that students have a general sense of academic self-efficacy (Chemers et al., 2001). This is a topic that remains a point of controversy among scholars who presently study student self-efficacy beliefs.

Based on his empirical research, Bandura (1997) asserted that self-efficacy beliefs are multidimensional, and that these beliefs, “should be measured in terms of particularized judgments of capability that may vary across realms of activity, under different levels of task demands, within a given activity domain, and under different situational circumstances” (p. 42). Other self-efficacy scholars agree. It has been found through various research studies that students with high self-efficacy beliefs for one domain of study may not have high self-efficacy beliefs across other academic domains (DiClemente, 1986; Hofstetter, Sallis, & Hovell, 1990; Pajares, 1996; Pajares & Miller, 1997). Usher and Pajares (2008) asserted that it is unreasonable to compare students’ domain specific self-efficacy judgments with their overall sense of academic self-efficacy. They noted that students who have high self-efficacy for one academic subject may have low self-efficacy for other subjects. Thus, a student’s overall sense of self-efficacy about his or her performance as a student, or academic self-efficacy, may be much lower than his or her self-efficacy for a particular academic subject. Although
many scholars study self-efficacy at domain, topic, or task specific levels (e.g., Pajares, 1996; Pajares & Miller, 1997; Pajares & Urdan, 2006), the idea of a general sense of self-efficacy continues to be studied and scales to measure these general beliefs continue to be created and refined (e.g., Chemers et al., 2001; Chen, Gully, & Eden, 2001).

Chen and colleagues (2001) argued that people have a general sense of self-efficacy that, “captures differences among individuals in their tendency to view themselves as capable of meeting task demands in a broad array of contexts” (p. 63). Through empirical research, it has been shown that a general sense of self-efficacy is positively related to other self-evaluation phenomena, including locus of control and self-esteem (Judge, Thoresen, Pucik, & Welbourne, 1999). General self-efficacy is also thought be positively related to the orientation of learning goals (Chen, Gully, Whiteman, & Kilcullen, 2000).

Chemers et al. (2001) studied the effects of general academic self-efficacy on students’ transition from high school to college. They believed that academic self-efficacy would affect both academic achievement and personal transition from the high school to the college environment. Chemers and colleagues investigated the relationship between overall academic self-efficacy and overall academic achievement and found these constructs to have a positive, statistically significant relationship. This supports the idea that students may hold general academic self-efficacy beliefs about their overall academic success.

Controversy remains regarding the level of specificity of student self-efficacy beliefs. I believe that students hold both overarching beliefs about their success as a student in general as well as specific self-efficacy beliefs that can differ by domain, topic,
and task. The following review of empirical research includes studies that considered self-efficacy beliefs at the academic, domain specific, and topic or task specific levels.

**Contemporary Empirical Self-efficacy Research**

Numerous research studies have been conducted concerning self-efficacy beliefs. Self-efficacy studies have been carried out across various domains, including psychology, medicine, sociology, and education. While providing further evidence of the positive relationship between self-efficacy beliefs and academic achievement may not have been the primary goal of the researchers conducting these studies, each showed that self-efficacy beliefs are positively and statistically significantly related to academic achievement.

**Self-efficacy beliefs in college students.** Chemers and colleagues (2001) conducted a study to determine if general academic self-efficacy beliefs had an effect on the transition from high school to college and on first-year college student academic performance. Chemers et al. (2001) hypothesized that academic self-efficacy would have a, “profound impact on the academic performance and personal adjustment of first-year college students as they navigate the demanding environment of university life” (p. 57). The sample for this study included 256 first-year college students.

Chemers and colleagues (2001) distributed survey packets to the students in order to gather self-reported data on perceived academic self-efficacy, social self-efficacy, optimism, current and anticipated academic performance, current and expected social adjustment, social support, expected levels of and the ability to cope with pressures and demands, and stress level. These surveys were administered at the beginning and end of
the winter quarter. Each participant’s high school GPA and final statistics course grade for the winter quarter were also obtained.

As expected, academic self-efficacy beliefs were found to be positively statistically significantly related to academic performance for first-year college students (Chemers et al., 2001). Chemers and colleagues (2001) concluded that, “students who enter college with confidence in their ability to perform well academically do perform significantly better than do less confident students” (p. 61). Academic self-efficacy beliefs were also found to be statistically significantly related to adjustment to college, mediated by a challenge-threat variable which was derived from the responses to the questions regarding expected levels of and the ability to cope with pressures and demands. From this evidence, Chemers and colleagues surmised that students with high levels of confidence and optimism regarding expectations about and the ability to cope with the college transition viewed the college experience as a challenge to successfully overcome, as opposed to a threat to be feared. Overall, first-year college students coming to college with higher levels of academic self-efficacy fared better during the transition from high school, both academically and personally.

Other empirical studies have produced similar results. Vuong, Brown-Welty, and Tracz (2010) investigated the relationship between academic success and academic self-efficacy beliefs in first generation college students. They surveyed 1291 college sophomores using the College Self-efficacy Inventory (CSEI). The CSEI contains questions regarding students’ demographics, intent to remain in school, GPA, and overall academic self-efficacy beliefs. Current GPA was used as the measure of academic achievement. Similar to the findings of Chemers and colleagues (2001), Vuong et al.
(2010) found that current GPA was positively statistically significantly related to general academic self-efficacy beliefs.

While current GPA is an indicator of academic achievement, as used in the study conducted by Vuong and colleagues (2010), it is also related to current academic standing. Hsieh, Sullivan, and Guerra (2007) conducted a study that examined the relationship between college student attrition due to underachievement and student self-efficacy beliefs. They found that students in good academic standing, those with a GPA of 2.0 or greater, had higher academic self-efficacy beliefs than did students on academic probation with a GPA of less than 2.0.

For students at various grade levels, it has been shown that academic achievement and, therefore academic standing, are positively related to student self-efficacy beliefs (Chemers et al., 2001; Hsieh et al., 2007; Pajares, 1996; Pajares, 2006; Pajares & Urdan, 2006; Vuong et al., 2010). The scholars who conducted the studies presented in this section considered general academic self-efficacy beliefs in relation to academic achievement. Below, I review the literature that explores the relationship between academic achievement and student self-efficacy beliefs related to the specific domain of statistics.

**Statistics self-efficacy.** The literature regarding student self-efficacy beliefs in the domain of statistics is scant at best. Until recently, scholars who studied various phenomena within the field of statistics often relied on the results of previous mathematics studies to support their research hypotheses (Garfield, 1995). Increasingly, scholars in the fields of statistics and statistics education have come to realize that statistics is a domain that presents its own set of challenges and difficulties, and that
mathematics simply cannot be used as a proxy any longer (Ben-Zvi & Garfield, 2005). One of the few studies that explored the self-efficacy beliefs of college statistics students was also an attempt to compose survey instruments for this domain that would produce valid and reliable inferences from data (Finney & Schraw, 2003). Based on Bandura’s (1997) theory of social-cognitive learning and definition of self-efficacy, Finney and Schraw (2003) developed two self-efficacy measures: a measure of current statistics self-efficacy (CSSE) and a measure of self-efficacy to learn statistics (SELS). Finney and Schraw (2003) defined current statistics self-efficacy as the, “confidence in one’s abilities to solve specific tasks related to statistics” (p. 164) and self-efficacy to learn statistics as the, “confidence in one’s abilities to learn the skills necessary to solve specific tasks related to statistics” (p. 164). Administering the CSSE to college students both at the beginning and end of a typical semester, Finney and Schraw were able to determine if self-efficacy for statistics increased over the period of a semester long introductory undergraduate statistics course. Using the SELS, administered at the same time as the CSSE, along with a self-efficacy for mathematics survey instrument, Finney and Schraw were able to compare self-efficacy to learn statistics with self-efficacy to learn mathematics.

Data for this study were collected from 154 undergraduate students taking an introductory statistics course offered through the Educational Psychology department of a large university. Besides the self-efficacy data collected, Finney and Schraw (2003) gathered data regarding previous statistics coursework, level of test anxiety, attitude toward statistics, mathematics self-efficacy, and final grade in the statistics course in which the participants were enrolled. The data gathered using the CSSE showed that the
college students’ current self-efficacy beliefs increased, on average, by nearly two standard deviations over the course of the semester (Finney & Schraw, 2003). A comparison between the data produced from the SELS instrument and the mathematics self-efficacy data showed that high self-efficacy beliefs for topics in the domain of mathematics did not necessarily imply high self-efficacy for statistical topics. This finding added to the growing evidence that mathematics cannot be used as a proxy for statistics (Ben-Zvi & Garfield, 2005; Garfield, 1995; Garfield & Ahlgren, 1988).

Hall and Vance (2010) conducted a study in which they investigated the relationship between peer feedback regarding students’ problem solving methods and student self-efficacy beliefs. The sample included 142 college students enrolled in a business statistics course. Participants were randomly assigned to either a feedback or no feedback group. All participating students completed a self-efficacy survey prior to the problem solving activity. The survey included questions about the students’ self-efficacy for statistics, level of confidence in their peers to provide effective feedback on their problem solving methods and skills, and level of confidence in their teacher to correct their methods of problem solving. The students in the no feedback group then completed three statistics problems independently. The participants in the feedback group were randomly separated into sets of three students. Each triad engaged in discussion via an online system. During the online discussions, group members provided information about their individual problem solving process, as well as feedback on their peers’ methods of solving the three problems. After all students had completed the three problems, participants completed the same self-efficacy survey that had been administered prior to the problem solving activity.
Hall and Vance (2010) found that, for students in the group that received peer feedback, a statistically significant positive relationship existed between scores on the three problems and post activity self-efficacy scores. For students in the no feedback group, this relationship was slightly negative but was not statistically significant. Hall and Vance posited that the feedback environment allowed students to observe peers and learn from peer models who may have been more capable of solving the statistics problems, thus positively affecting self-efficacy beliefs and achievement. This finding is directly related to Bandura’s (1995) idea that vicarious experience, or modeling, is one way to increase student self-efficacy beliefs.

**Self-efficacy beliefs and modeling.** According to Albert Bandura (1995), vicarious experiences (i.e., modeling), or observing a similar peer with high self-efficacy beliefs achieve at a high academic level, is one of four ways of positively influencing student self-efficacy beliefs. Students often learn more from their peers than from instructors because they are able to discuss problems in a common language that makes more sense to them than the vocabulary often used by teachers (Cohen, 1994b). When a student with lower self-efficacy beliefs is working in a group with peers of similar age, some of whom hold higher self-efficacy beliefs for the task or topic at hand, Cohen (1994b) noted that less efficacious students have the opportunity to observe, ask questions, discuss, argue, negotiate, explain and put forth unique ideas in a trusting environment among a small group of classmates. By observing and engaging with peers as a group successfully navigates a task, less efficacious students gain mastery experiences (Bandura, 1995) and acquire new skills, which in turn increase self-efficacy beliefs (Schunk, 1984) and promote academic achievement.
Schunk and Hanson (1985) studied the effects of peer observation on student self-efficacy beliefs and academic achievement. Working with 72 students between the ages of 8 and 10 years who, according to their classroom teachers, struggled with basic subtraction skills, Schunk and Hanson began by administering a pretest that focused on self-efficacy beliefs for solving subtraction problems. Following the self-efficacy pretest, the students were given a subtraction skills pretest. After the two pretests were complete, the students were randomly assigned to one of six groups: male mastery model, male coping model, female mastery model, female coping model, teacher model, or no model. According to Schunk and Hanson, mastery models were model students, presented via videotape, who appeared very confident and fearless in their approach to the task at hand. Coping models, on the other hand, were model students, presented via videotape, who took a more gradual, somewhat apprehensive approach to completing the task. Teacher models were simply teachers, with no peers present. None of the models in the videotapes were members of the classes from which the participants were selected.

Students in each of the six groups were placed in sets of four or five students each. Each of the groups assigned to observe one of the four peer models was shown a 15 minute video that presented a non-classmate peer learning subtraction skills. The non-classmate peer in the video was the same sex as the observing students. Schunk and Hanson (1985) posited that, at the time of experimentation, students might have reacted more favorably and paid more attention to models of the same sex as themselves. Those groups assigned to observe the teacher model watched a 15 minute video of a teacher presenting methods of solving subtraction problems, with no one but the teacher present in the video. Finally, the students assigned to the no model category received typical
teacher led instruction that was presented to all students in the study and these students observed no model.

Following all treatments and instruction, the students were administered a subtraction self-efficacy posttest and a subtraction skills posttest. The self-efficacy posttest was identical to the pretest and the subtraction skills posttest was a parallel form of the pretest. The parallel form was used to ensure that students did not improve due simply to familiarity with the problem set.

Schunk and Hanson (1985) found that the self-efficacy beliefs of students in the peer model groups increased statistically significantly more than those in either the teacher model or no model groups. There was no significant difference, however, between the peer mastery model and coping model groups. That is, it did not seem to matter if students observed their peers easily completing a task or if they watched their peers struggle while learning; as long as students observed their peers achieving success, the observing students’ self-efficacy for the specific task increased more than if they had observed a teacher or no one at all.

**Conclusion**

The contemporary research literature contains overwhelming evidence that increased self-efficacy beliefs can help promote academic achievement for students at all academic levels, including college. It is clear from Albert Bandura’s (1997) theoretical work regarding self-efficacy, as well as the empirical literature, that one of the ways to increase student self-efficacy is through modeling, or the observation of a similar peer who holds high self-efficacy beliefs and achieves at a high academic level. The Schunk and Hanson (1985) study is one of few in the body of self-efficacy literature that truly
captured the role that modeling plays in positively influencing student self-efficacy beliefs. Further research similar to that of Schunk and Hanson is necessary to understand the practical methods educators can employ to help students increase their self-efficacy beliefs. In addition, research similar to the work of Pajares is needed to examine how to foster college students’ topic specific self-efficacy beliefs in the domain of statistics.

Statistics Education

Understanding new statistical ideas and topics is a challenge for many students (Garfield, 1995). Even at the undergraduate and graduate college levels, students in introductory courses struggle to comprehend various statistical concepts due to their unfamiliarity and unease with the content. For example, after completing an introductory statistics course, many students do not understand that a larger sample is more likely to produce statistical values that more accurately reflect the population than is a smaller sample (Zieffler et al., 2008).

There are myriad reasons that students have difficulty fully understanding statistical concepts. Many scholars in the field of statistics education, researchers and practitioners alike, believe that one key reason students struggle in the statistics classroom is that instructors do not focus on fostering students’ statistical literacy and statistical reasoning skills, instead choosing to teach the processes for simply solving problems and producing the correct answer (Ben-Zvi, 2005; Ben-Zvi & Garfield, 2005; delMas, Garfield, Ooms, & Chance, 2007; GAISE College Report, 2005; Garfield, 1995; Garfield, 2005). The GAISE College Report (2005) stated that, “the desired result of all introductory statistics courses is to produce statistically educated students, which means that students should develop statistical literacy and the ability to think statistically.
Achieving this knowledge will require learning some statistical techniques, but the specific techniques are not as important as the knowledge that comes from going through the process of learning them” (p. 5).

In a review of contemporary research concerning how students learn statistics, Garfield (1995) found that in statistics classrooms, the use of collaborative learning fostered improved productivity, better attitudes toward learning, and greater academic success. She also found that when small collaborative groups of students were engaged in activities, the students learned to effectively argue their ideas. These small group discussions helped students become more involved in their learning process, which in turn helped to create deeper understanding. Through her research in the field of cognition, Lovett (2001) has come to believe that collaboration among students in the classroom will further student statistical reasoning, literacy, and understanding.

Ben-Zvi and Garfield (2005) found that one of the major benefits of using collaborative group work in the statistics classroom is that students are given the opportunity to communicate using statistical language. Similar to a foreign language class, students practice using new terminology in a way that makes sense to them and their peers. When communicating with their peers in small, collaborative groups, students who struggle with statistical language may emulate their higher performing peers, which can lead to an increase in their self-efficacy for statistical language, thus helping them to understand and academically achieve (Bandura, 1995; Garfield, 1995). This collaborative group interaction helps students understand whether or not they are using the vocabulary of statistics accurately. During these collaborative discussions, the instructor can wander among the groups, listening to ensure that students accurately
comprehend the material. The ability to understand and correctly use statistical terminology is a key skill necessary for students to become statistically literate (Utts, 2003).

Working together with peers in small groups, discussing topics with peers, hearing various perspectives, and learning from peer models, all of which are aspects of collaborative learning, are ways of promoting academic achievement across various domains, but particularly in statistics (Cohen, 1994b). A collaborative learning environment provides students with space to discuss their ideas in a small, safe setting where they feel comfortable putting forth and negotiating their ideas (Johnson & Johnson, 1989). According to the literature regarding statistics education, while collaborative learning has shown great promise as a method of instruction, it is not widely employed by college statistics instructors (Ben-Zvi & Garfield, 2005; Garfield, 1995, 2005). In response to the recommendation put forth in the GAISE College Report (2005), I hypothesized that collaborative learning can increase the self-efficacy beliefs of statistics students and, thus, can be the process by which students develop statistical literacy and thinking and improve their performance in the statistics classroom.

It is important to examine collaborative learning and students’ self-efficacy beliefs within a specific domain. The domain of statistics is uniquely challenging (Ben-Zvi & Garfield, 2005; Garfield, 1995; Garfield & Ahlgren, 1988), yet collaborative learning has been advanced as a way to help students navigate these challenges (Cohen, 1994a; Garfield, 1995; Johnson & Johnson, 1989). The often complex and abstract topics associated with statistical sampling can be particularly difficult for introductory statistics students to understand (Fecso et al., 1996; Gelman & Nolan, 2002; Nguyen, 2005; Utts & Utts, 2003).
Heckard, 2006; Yilmaz, 1996). Therefore, I chose to conduct the present study with students studying statistical sampling to determine if the relationship between collaborative learning and students’ self-efficacy beliefs can be leveraged to foster better understanding of the topics associated with statistical sampling.

**Teaching and Learning Statistical Sampling**

Statistical sampling, or survey sampling, is a topic typically covered in introductory undergraduate statistics courses or taught as a semester-long course unto itself. The topics associated with statistical sampling, such as data collection, sample size determination, research design, understanding and developing sampling methods, margin of error and bias, and survey development and implementation can be difficult to comprehend for students new to this material (Fecso et al., 1996; Gelman & Nolan, 2002; Nguyen, 2005; Utts & Heckard, 2006; Yilmaz, 1996). A review of the research literature revealed several problems undergraduate students encounter when learning about statistical sampling.

Undergraduate statistics students often find the topics associated with statistical sampling to be overly abstract and complex (Nguyen, 2005; Yilmaz, 1996), requiring analytical skills that can be difficult to teach (Yilmaz, 1996). Because introductory statistics students typically have diverse mathematical and statistical backgrounds, as well as varying academic expectations, instructors often struggle to teach the material associated with statistical sampling (Fecso et al., 1996; Yilmaz, 1996). For many undergraduate students, a course in statistical sampling or an introductory course in which statistical sampling is a topic covered over a period of a few weeks may to be their only exposure to the subject. This ‘once-in-a-lifetime’ experience for students who will
not become experts in some quantitative field of study creates a quandary for instructors (Fecso et al., 1996; Yilmaz, 1996). Instructors must decide whether to delve deeply into a few statistical sampling topics, or to cover the material more broadly, offering students a brief survey of many topics. Because this is often the first time many undergraduate students will discuss the topics associated with statistical sampling, and the time allotted to cover these topics is often brief, scholars and instructors have suggested that these topics be taught in a conceptual manner as opposed to a theoretical, deeply mathematical way (Fecso et al., 1996; Utts & Heckard, 2006; Yilmaz, 1996).

Beyond these general issues students often face when learning about statistical sampling, research scholars have also established that students encounter various specific issues when learning this new topic. In a study with high school statistics students, Saldanha and Thompson (2002) found that students have difficulty understanding the difference between the idea of a sample and the concept of repeated sampling, and what each means in relation to the population. They noted that, in practice, researchers typically make inferences about the population based upon a single sample, a process that is accepted as sufficient in the real world. They believed, however, that students should understand why researchers accept this single sample process as sufficient and that teaching students about repeated sampling can help them understand the reasoning behind the practice. In a statistics course focused upon basic, conceptual understanding, however, these reasons behind various practices are often not addressed (see Utts & Heckard, 2006).

Schwartz and Goldman (1996) conducted a study with college sophomores to gain insight into their process of developing a sampling plan. They asked the 24 students
to design a small study to investigate the relationship between eating fast food and being overweight. Nineteen of the 24 students suggested that a survey should be conducted only at fast food establishments, ignoring people who do not eat at fast food restaurants. Schwartz and Goldman noted that this short sightedness is common when students are beginning to learn how to develop sound sampling plans.

Watson and Moritz (2000) studied students’ understanding of the concepts of sample and sampling. Students enrolled in various mathematics courses were surveyed and participated in follow up interviews. Watson and Moritz found that students’ understanding of sampling varied greatly, ranging from a basic understanding of the word sample but inability to apply understanding to various contexts, to a fully developed understanding of sample, sampling methods, and sampling bias and the ability to apply this knowledge to various situations. This wide range of prior knowledge can present difficulties for instructors teaching these concepts. When beginning the lessons associated with the concepts of sample and sampling, Watson and Moritz recommended teaching through examples and scenarios that are relevant to the students in each class in order to assess students’ levels of knowledge. For example, when teaching high school students, they suggested using voluntary response surveys published in popular magazines as a starting point for discussions regarding sample and population. They believed that a familiar, relevant starting point can help students exhibit their understanding without being overwhelmed by the statistical concepts.

One instructional method that scholars and college instructors have found to help students overcome various misconceptions and to understand these complex topics related to statistical sampling is to provide students with hands-on, real world projects.
and activities (Chang, Lohr, & McLaren, 1992; Fecso et al., 1996; Gelman & Nolan, 2002; Hodgson & Burke, 2000; Kelly, 2010; Mills, 2002; Richardson, 2003; Warton, 2007; Yilmaz, 1996). Kalsbeek (Fecso et al., 1996) believed that the most important strategic decision to make when developing a course or instructional unit regarding statistical sampling “is to make the student’s experience as close to hands-on as the classroom setting will allow. This exposure, as mounting evidence seems to support, is critical to explain, reinforce, or simply make real some of the concepts being taught” (p. 333). Using real world, active projects can help increase undergraduate students’ academic achievement, course satisfaction, and student confidence (Kelly, 2010), makes the abstract topic of statistical sampling more accessible for undergraduate students (Fecso et al., 1996), and can be more interesting for instructors than a typical course structured around lectures and exams (Chang et al., 1992; Kelly 2010). Offering hands-on activities when teaching statistical sampling also helps to emphasize to students the importance of understanding ‘the method’ rather than the quest for ‘the answer’ (Warton, 2007), which can help create statistical literacy (Ylimaz, 1996), as recommended by the GAISE College Report (2005).

One way in which undergraduate instructors have offered hands-on, real world statistical sampling activities to their students was by assigning the class one large project (Gelman & Nolan, 2002; Kelly, 2010). At the beginning of the semester or instructional unit on statistical sampling, instructors collaborated with their students to develop a sampling project idea. Once the class and instructor agreed upon an idea, the class was divided into small working groups, each of which was responsible for a portion of the project, including development of the sampling plan, survey design, data collection, data
analysis, and final report writing. Some of the small, collaborative group work occurred in class and some outside of class. Although each group focused primarily upon one aspect of the project, all student groups provided regular, in-class status reports for their peers. During these report sessions, students received handouts to help guide the whole class, collaborative discussions and both students and the instructor provided constructive feedback to each of the groups.

Gelman and Nolan (2002) noted that this whole class, collaborative project had several advantages over the typical lecture-based course or instructional unit on statistical sampling. In their courses, students were able to choose to work on an aspect of the project that best aligned with their skills and interests. This helped to keep students interested and motivated while working on the project. Working as a class on one large project created a feeling of collective effort on the part of the students, and thus students were more interested and invested in completing the project. Because the work was distributed, student workload was manageable, allowing the students to focus upon fully understanding the material as opposed to rushing to finish their work. The students were excited about the project and some even followed up after completion of the course to work with the data in other ways (e.g., honors theses, further research studies, etc.).

Kelly (2010) found that, compared to students who received primarily lecture-based instruction, the students who participated in the whole class collaborative project exhibited higher academic achievement and were more satisfied with their experience in the class. She made reference to the amount of time it took for students to become comfortable with this type of learning environment, stating that, “the students, while initially reluctant to speak in class and wanting me to make decisions, gradually, [and]
with encouragement, increased in confidence and quite soon took over complete responsibility in running the survey” (Kelly, 2010, p. 463). This increase in student confidence can lead to increased self-efficacy beliefs and, in turn, promote higher academic achievement for students studying statistical sampling (Bandura, 1995). Kelly (2010) also stated that this type of hands-on learning environment was more interesting for both the students and the instructor.

Hodgson and Burke (2000) echoed the sentiments mentioned above regarding the benefits of active, hands-on, real world tasks for students learning statistical sampling. They offered some recommendations to help instructors implement this type of learning environment and gain awareness of their students’ levels of knowledge and understanding. The authors suggested that conducting informal, formative assessment is crucial during the active learning tasks. They noted that this type of observation and assessment can help highlight areas where students are confused, and allow instructors to make necessary adjustments to future instruction and activities. Hodgson and Burke also recommended a period of ‘debriefing’ at the end of each class. They believed that allowing students to share and discuss their final thoughts and questions helped all students gain an accurate, deep understanding of the concepts related to statistical sampling.

Undergraduate students studying statistical sampling for the first time often have difficulty comprehending the rather abstract and complex topics associated with this statistical topic (Fecso et al., 1996; Gelman & Nolan, 2002; Nguyen, 2005; Utts & Heckard, 2006; Yilmaz, 1996). The recommendations from various scholars and instructors to use real world, hands-on activities to teach statistical sampling aligns with
the theory and practice of collaborative learning (Johnson & Johnson, 1989). Small, student led collaborative groups working together on real world, hands-on tasks is an instructional method that can help undergraduate statistics students gain a deep understanding of the topics associated with statistical sampling and can help to increase student satisfaction and confidence (Chang et al., 1992; Fecso et al., 1996; Gelman & Nolan, 2002; Hodgson & Burke, 2000; Johnson & Johnson, 1989; Kelly, 2010; Mills, 2002; Richardson, 2003; Warton, 2007; Yilmaz, 1996), leading to increased self-efficacy and, thus, greater academic achievement (Bandura, 1995). However, there is no research that establishes students’ self-efficacy beliefs as the mechanism through which collaborative learning affects students’ understanding of the concepts associated with statistical sampling. The present study was an attempt to fill this gap in the research literature. In the following section, I present a review of the collaborative learning literature.

**Collaborative Learning**

**Defining Collaborative Learning**

Teaching students to work collaboratively with their peers within a classroom setting can help promote high academic achievement for students at all achievement levels (Cohen, 1994a; Johnson & Johnson, 1989). Cohen (1994b), defined group work as, “students working together in a group small enough that everyone can participate to carry out their task without direct and immediate supervision of the teacher” (pp. 1-2). Collaborative learning is a form of group work where students openly discuss their thoughts and ideas, productively argue, and negotiate with one another as they work together to come to an agreed upon solution to the problem at hand (Gerlach, 1994;
Palinscar and Herrenkohl, 1999). Collaborative group work is often confused with cooperative group work, and some scholars use the terms interchangeably. But Palinscar and Herrenkohl (1999) pointed out that, while collaboration requires cooperation, not all cooperative learning requires true collaboration. They believed the goal of collaborative group work is for the group to reach a common end, whether that is a common idea, one group project, or a combined performance that all members of the group agree is the best representation of their final product.

The Benefits of Collaborative Learning

In the recent past, many teachers viewed collaborative learning as a passing fad that was more of a burden than a helpful instructional tool (Gerlach, 1994). Now, however, the tide seems to have turned. Gerlach (1994) found that, “listening to different points of view about how to solve problems or to different perspectives on issues helps students to reach deeper levels of understanding about their subjects” (p. 10). These collaborative discussions, which often incite conceptual change, can in turn increase student achievement (Slavin, 1990). It has been shown that collaborative learning is a more successful method of promoting academic achievement than individual or competitive learning (Johnson, Maruyama, Johnson, Nelson, & Skon, 1981).

When higher-order understanding and thinking is the goal, collaborative learning is often a good mechanism to help achieve this (Cohen, 1994a). The act of engaging in scholarly discussion with peers is often the catalyst for achieving this level of understanding. Explaining a concept to a peer or group of peers who may not have a firm understanding helps the student who is providing the explanation to solidify his or her conceptual knowledge of the topic (Chang & Zellner, 2011; Cohen, 1994a, 1994b;
Johnson & Johnson, 1989, 2002, 2009; Slavin, 1990). At the same time, the student or students who are listening to their peer and are asking questions about the explanation are refining their ideas about the topic until they are comfortable that they have an accurate understanding. Kuh (2008) noted that, “collaborative learning combines two key goals: learning to work and solve problems in the company of others, and sharpening one’s own understanding by listening seriously to the insights of others, especially those with different backgrounds and life experiences. Approaches range from study groups within a course, to team-based assignments and writing, to cooperative projects and research” (p. 1). It is through listening to, observing, and emulating their higher achieving peers that members of a collaborative group can increase their self-efficacy beliefs and, in turn, achieve academic success (Bandura, 1995).

As college students graduate and enter the workforce, they will likely enter a profession in which they will be expected to interact with and often collaborate with others. Small collaborative group work during college can be a training ground during which students practice and hone the appropriate social skills that help them become effective members of the workplace (Cohen, 1994b). Often, students who do not interact during whole class discussion are more comfortable participating in a small group setting. When a group works collaboratively, students within the group often become friendly with one another and begin trust each other (Deutsch, 1968). Allowing students to make decisions for their group and to work together to create a common outcome provides them with the opportunity to plan how they will complete the task and gives them control over their learning environment (Cohen, 1994a).
Social Interdependence Theory

The social interdependence of groups is characterized by the idea that each individual member of a group is affected by the actions and behaviors of the other members of the group (Johnson & Johnson, 1989). When working effectively, a collaborative learning environment is distinguished by the presence of positive social interdependence, which results in promotive interaction (Johnson & Johnson, 1989). Usher and Pajares (2008) asserted, “social models play a powerful role in the development of self-efficacy, especially when students are uncertain about their abilities or have limited experience with the academic task at hand” (p. 753). According to social interdependence theory, peer modeling is one of the mechanisms by which learning and achievement occur (Johnson & Johnson, 1989). Collaborative learning both depends on and promotes promotive interaction, the key characteristic of positive social interdependence. Social interdependence has served as the theoretical basis for the large majority of empirical research studies conducted regarding collaborative learning (Johnson & Johnson, 2002).

In order to achieve full positive social interdependence, a collaborative group must be structured in both an outcome interdependence and means interdependence manner (Johnson & Johnson, 1989). If the group structure is based solely on outcome interdependence, in the absence of means interdependence, Johnson and Johnson (1989) claimed this will result in “co-action or nonspecific/unstructured interaction toward mutual outcomes” (p. 25). That is, members will either work interactively with no specific role or task assignments (i.e., co-action), or members will work independently (i.e., nonspecific/unstructured interaction) and then sum their individual efforts to create
the whole. This group structure does not allow for collaboration because no negotiation, constructive argumentation, or discussion takes place to encourage the creation of a common, agreed upon final product.

Outcome interdependence is divided into two categories: positive goal interdependence and positive reward interdependence (Johnson & Johnson, 1989). Positive goal interdependence takes place when individual group members believe they are able to reach their goal only when other group members with whom they are linked also reach their goals. Positive reward interdependence occurs when all group members receive the same reward for completing the task successfully.

If a group is based on a means interdependence structure with no outcome interdependence structure present, then members of the group are working to attain independent goals. Johnson and Johnson (1989) said, in this scenario, members help one another achieve individual goals that are unrelated to one another. Again, this type of group structure is not collaborative in nature because group members work toward individual goals with no regard for the overall group goal. Thus, as Johnson and Johnson noted, in order for a collaborative group to work effectively, the group structure must include both outcome and means interdependence.

When group members believe two or more members are required to coordinate efforts in order to successfully complete a task, Johnson and Johnson (1989) noted that the group is structured in a means interdependence fashion. There are three types of means interdependence: resource interdependence, role interdependence, and task interdependence. Resource interdependence is when each individual in the group has only one piece of the information or materials necessary to complete a task. Members
then have to pool their resources in order to complete the task. Role interdependence is when group members are assigned specific roles (e.g., group leader, scribe, reporter, etc.) that are essential for the group to function. These roles are interdependent upon one another and group members must work together to complete the task. Task interdependence is when each group member is assigned a different, yet interrelated task. Therefore, one group member must complete his or her task in order for another group member to start and complete the next task.

In order for true collaboration to occur, both outcome interdependence and means interdependence group structures must exist (Johnson & Johnson, 1989). When both structures are present within a group, positive social interdependence and, hence, promotive interaction, can take place. Positive social interdependence is the key mechanism that fosters collaborative learning. When small student led groups work together in a collaborative manner, students can perform at higher academic levels (Cohen, 1994b).

Research has shown that collaborative learning, and hence positive social interdependence, promotes higher academic achievement than competitive or individual learning (Johnson & Johnson, 1989). Johnson and Johnson conducted a meta-analysis of 189 research studies that investigated cooperative or collaborative learning. Regarding the promotion of academic productivity and achievement, they found that greater than 50% of the studies favored a cooperative or collaborative learning environment over a competitive or individualistic setting. Johnson and Johnson also found that the students who engaged in cooperative or collaborative learning employed higher quality reasoning skills more often than students who worked competitively or alone. In the large majority
of the studies it was not the case that the lower achieving students within the collaborative learning groups were simply ignored or told the correct answers. Instead, students actually taught one another and group discussion was the primary catalyst for higher-order thinking for all students in the groups. These acts of peer tutoring and discussion are ways in which students act as models for one another during collaborative learning activities (Hancock, 2004).

**Collaborative Learning and Academic Achievement**

The use of collaborative learning in the classroom advances academic achievement to higher levels than are attainable through individual or competitive learning (Gerlach, 1994; Johnson & Johnson, 1989). Emerson, Phillips, Hunt, and Alexander (1994) presented case studies that offer support for this. In one of these studies, a college level mathematics instructor who had taught using traditional lecture methods for almost 20 years before instituting collaborative learning in the classroom found that, “teaching collaboratively has enabled me to get through more material, and students have achieved a deeper understanding, worked harder, and enjoyed it” (Emerson et al., 1994, p. 83).

**Collaborative learning in university classrooms.** Until recently, collaborative learning was primarily considered an instructional method for use in elementary and middle-schools, with little collaborative learning occurring in high school and postsecondary classrooms (Johnson & Johnson, 1989). Of late, however, college and university instructors have begun to recognize the benefits collaborative learning can have for their students and there has been an increase of collaborative learning activity in college classes across various disciplines. Emerson and colleagues (1994) presented a
case study conducted by one of the authors (Emerson) in his college anatomy and physiology course. In this study, Emerson replaced lecture style instruction with collaborative, student led, small group learning.

For this study, each class began which a brief introductory whole class discussion and explanation of present class objectives followed by a specific collaborative small group activity, during which students had to analyze the task, evaluate information, and come to a common understanding of the task at hand (Emerson et al., 1994). Each class ended with a full class summarization discussion and presentation of connections to other topics discussed earlier in the semester. At the end of the semester, Emerson found that a collaborative learning environment produced higher end of course grades than did a lecture style course and that students enjoyed the collaborative group work more than lecture. In fact, the results of the case study showed that, “more than half the students thought that the activities were more interesting than those used in other methods and that they had helped them to learn the material better than other methods. Seventy-two percent thought the activities should be used in the next semester” (Emerson et al., p. 90).

A study sponsored by the Association of American Colleges and Universities found that increased student collaborations, both inside and out of the classroom, promoted better student understanding and problem solving skills and increased academic achievement (Kuh, 2008). Grover (2010) witnessed this at work in her college science courses, noting that collaborative learning both inside and out of her classroom helped her students produce higher quality work. She used collaborative learning in all of her classes, which include organic chemistry, protein biochemistry, and nucleic acids
biochemistry, among others. Grover offered the following examples of successful collaborative learning from her personal experiences.

The introductory organic chemistry course Grover (2010) presented in the first example included 24 undergraduate students. She noted that most of her colleagues taught this course using more traditional, lecture-based methods of instruction. Grover, however, often employed the use of worksheets that were to be completed collaboratively in small student groups. These worksheets included tasks such as creating a list of key topics generated from the previous night’s reading. Grover had the student groups limit the number of topics they considered ‘key’, therefore forcing the students to negotiate among themselves as to whether or not a topic warranted inclusion on the list. She listed several benefits of having students work in small collaborative groups in this course. She found that students are more comfortable asking questions of their peers than seeking help from the instructor. Small groups also provided a forum for insightful and thoughtful discussion where multiple perspectives were provided by the various students in each group. Grover also found that both student interest and academic performance were higher when students worked collaboratively within a small group of their peers than when they worked alone.

The two biochemistry course examples Grover (2010) presented were more advanced, thus she was comfortable using collaborative group activities that were more complex than the worksheets used in her introductory organic chemistry course. In the protein biochemistry course, students were separated into the same groups of three or four for the duration of the course. Each group had the freedom to design and test some of their own laboratory projects. Grover found that this level of autonomy helped her
students develop stronger beliefs in their ability to learn and fully understand the material, thereby increasing their self-efficacy for the topic at hand (Bandura, 1977). Grover (2010) also found that students attended to the textbook material more acutely when they recognized that a full understanding of the information was necessary to help in creating their own laboratory protocols. Grover (2010) noted that students gained a better understanding of the material when using a collaborative learning approach and, “the depth of student learning is much greater in this format than when they are lectured on the same material” (p. 326). It stands to reason that this deeper understanding and learning would likely promote higher academic performance.

**Collaborative learning and argumentation.** Discussion is one of the key mechanisms through which students help one another gain a better understanding of ideas and topics when working together in collaborative groups (Cohen 1994b). Chinn (2006) asserted that argumentation is a productive tool employed by students during small group discussion. According to Chinn, there are four main benefits of argumentation. These include developing a deeper understanding of course content, increased student interest and motivation, better student performance on problem solving tasks, and students’ increased ability to effectively argue. Chinn stated that the inherent interest created by the controversy of argumentation can increase students’ intrinsic motivation. He also believed that arguing with peers within a small collaborative group allows students more freedom to express their thoughts and ideas than in a traditional classroom setting. He warned that argumentation can become aggressive and counter-productive to the task at hand. Therefore, instructors must monitor group arguments to ensure they are productive and effective. Erkens and colleagues (2006) noted that, “an important distinction needs
to be made here between argumentation for reaching agreement and argumentation for convincing the partner [or group] of one’s own opinion, or even enforcing one’s beliefs” (p. 260). Hence, Erkens and colleagues (2006) suggested that instructors discuss this distinction with students prior to small group work in order for effective collaboration to occur.

Because collaborative group argumentation can increase student interest for the task at hand, students working in collaborative groups can become more captivated by the activity than students working independently. Chinn (2006) found that, “argumentation was marked by higher levels of engagement than was the discourse of traditional recitations” (p. 357). This increased interest can lead to a deeper understanding of the course material and can promote higher academic achievement (Bandura & Schunk, 1981; Rotgans & Schmidt, 2011).

**Computer-supported collaborative writing.** Chinn (2006) noted that argumentation that occurs during small collaborative group discussion is a mechanism by which students develop greater interest in the topic at hand. Erkens and colleagues (2006) offered support for this claim through their research with students working collaboratively in a computer-supported environment. Their COSAR project considered the depth of interactions between students working collaboratively as well as the problem solving skills, understanding, and academic achievement of these students as they collaboratively constructed essays while working in a computer-supported environment.

Erkens et al. (2006) developed a study designed for 16 to 18 year old students to work together writing argumentative essays. The students chose a position of either support or opposition of a current topic and were tasked to write a compelling argument
to bolster their position. Each student in the study was paired with a peer and the dyads were instructed to use a computer program called TC3 to help them collaborate on this assignment. TC3 provided students with several web-based tools, including a shared text editor, an online chat tool, various features that allowed students to create and organize their ideas, and access to resources that were intended to help foster online collaboration. Each student was expected to collaborate with his or her partner, but no intervention from the instructor or researchers took place if collaboration did not occur within each dyad.

From the results of this study, Erkens and colleagues (2006) found that, “coordination and discussion of specific content of goals, knowledge, and formulation positively influenced the argumentative quality of the final product” (p. 233). In other words, those students who used the online tools effectively and worked together to discuss their ideas with one another during the collaborative learning process produced higher quality work and attained higher academic success than pairs of students who chose not to collaborate. From the results of this study, Erkens and colleagues (2006) surmised that, “shared knowledge construction is an essential part of collaborative learning – both on a metacognitive level and on a specific content level” (p. 233).

**Social interdependence, collaborative learning, and academic achievement.**

Social interdependence has served as the theoretical basis for the majority of research studies focused on collaborative learning (Johnson & Johnson, 2002). The idea behind social interdependence is that, when students are working together in collaborative groups, the outcome for each individual member of the group is affected by the actions and behaviors of the other members of the group (Johnson & Johnson, 1989). When small classroom based groups work in a manner that creates positive social
interdependence, which results in promotive interaction, students gain a deeper understanding of the topic and achieve at higher levels.

It is this idea of creating and fostering promotive interaction that in turn promotes greater academic achievement. Johnson and Johnson (1989) expounded on this by stating:

Promotive interaction may be defined as individuals encouraging and facilitating each other’s efforts to achieve, complete tasks, and produce in order to reach the group’s goals. It is characterized by mutual help and assistance, mutual influence, the exchange of needed resources, interpersonal feedback, intellectual challenge and disagreement, the advocacy of committed efforts to achieve, and lower anxiety about performance. It is through promotive interaction that collaborators get to know each other as persons. (p. 29)

Promotive interaction leads to and is fostered by positive social interdependence, which in turn leads to more effective collaboration and higher academic achievement than is attainable in competitive or individual learning environments.

Chang and Zellner (2011) examined the effects of positive social interdependence on student achievement and student attitude. This research study included 144 undergraduate participants from three different universities. The students were enrolled in three separate college courses, including two teaching methods courses and a human resource development course. The students at each university were randomly assigned to three groups: a positive interdependence group, a group processing group, and a no interaction group. Group processing, according to the Chang and Zellner, is a process by which group members intermittently assess individual members’ contributions to the group and make the changes necessary to ensure that the group will successfully complete the task. The participants in the positive interdependence and group processing groups were further divided into small collaborative working groups. The participants in
these groups collaborated via an online environment, primarily utilizing online discussion.

Prior to the beginning of the semester, the instructors of the three courses were trained by the researchers on how to teach effective collaboration skills to their students (Chang & Zellner, 2011). The instructors were also provided with activities to use when instructing their students how to work in a collaborative learning environment. This training was necessary to encourage collaboration among the students in both the positive interdependence and group processing groups. At the beginning of the semester, the instructors of each of the three courses spent two days working face to face with their students in the positive interdependence and group processing groups teaching them how to foster effective collaboration. Following the two days of collaborative skills instruction, the students spent three weeks working on a specific assignment relevant to the class in which they were enrolled.

At the end of the three week period, Chang and Zellner (2011) found that using positive interdependence strategies was more effective in promoting higher academic achievement than was using group processing strategies or having no interaction with other students. This result supports the work and testaments of Johnson and Johnson (1989). Chang and Zellner’s (2011) findings indicated that positive social interdependence is a means of promoting academic achievement.

**Collaborative Learning and Modeling**

The literature on collaborative learning indicates that the foundation of effective collaboration is positive interaction among group members (Johnson & Johnson, 1989). This level of collaboration often requires students to learn and practice new social skills
(Cohen, 1994b; Slavin, 1990). In a meta-analysis of 189 studies regarding cooperative and collaborative learning, Johnson and Johnson (1989) found that student members of collaborative groups who were typically lower achievers were almost always included in group discussions and were made to feel like contributing members of the group. Higher achieving students modeled their understanding and ability to achieve for those students who were struggling with the material, and the discussions that came from these interactions served as a catalyst to promote higher-order thinking among all students in the group. The opportunity to observe a peer reach certain goals and achieve academically at high levels provides the support many lower achieving students need to positively affect their academic understanding and achievement (Chang & Zellner, 2011; Cohen, 1994b). Cohen (1994b) noted that, “when groups engage in cooperative tasks, they are more likely to form friendly ties, to trust each other, and to influence each other than when the task stimulates competition among members” (p. 17). Therefore, collaborative learning is superior to competitive or individual learning as a method of fostering an environment in which higher achieving students can act as positive models for and constructively influence their lower achieving peers’ academic accomplishments. Thus, collaborative learning may be a powerful instructional method of helping students bolster their self-efficacy beliefs and, in turn, increase their academic performance. However, further intervention research is needed to help establish a causal connection between collaborative learning and students’ self-efficacy beliefs.

**When Is Learning Collaborative?**

Given the literature reviewed in this section, it is important to identify when students are engaging in thoughtful discussion using appropriate academic language, are
arguing and negotiating productively to come to an agreed upon solution, and are supporting and encouraging one another to complete the task, all leading to high-quality collaborative learning. It can be difficult for instructors to determine whether true collaborative learning is occurring within small student groups (Palinscar and Herrenkohl, 1999). However, scholars have noted various specific student interactions that occur during collaborative learning activities. These indicators of quality collaboration include peer negotiations that use proper academic language (Dillenbourg, 1999; Grover, 2010) and engagement in deep, thoughtful discourse during which students are comfortable articulating their thoughts and ideas (Cohen, 1994b; Erkens et al., 2006; Grover, 2010; Ploetzner, Dillenbourg, Preier, & Traum, 1999). Another means to assess whether or not collaboration is occurring is to observe for periods of productive argumentation (Chinn, 2006; Dillenbourg, 1999; Erkens et al., 2006). In order to collaborate and come to an agreed upon final product, students must effectively argue their ideas and points of contention and come to resolution. These effective arguments, in turn, lead to deep understanding of academic topics (Chinn, 2006). Finally, when students working together in small groups are supportive of one another, encourage group members to succeed, and help one another reach the overall group goal, the group experiences positive social interdependence, the underlying phenomenon that fosters collaborative learning (Johnson & Johnson, 1989).

**Intervention Research**

Intervention research has been used to conduct studies in the field of education for decades (Hsieh et al., 2005). Levin and O’Donnell (1999) defined an intervention as a process in which the researcher or teacher intervenes to compare an alternative method.
with a more commonly practiced method. The results of implementing the alternative method are then evaluated and recommendations are made based upon the resulting analyses. During the 1970’s and 1980’s, the number of intervention research studies in education increased dramatically after the publication of what is thought to be the seminal work on educational intervention research written by Campbell and Stanley (1963). While other scholars have written more contemporary works regarding quasi-experimental studies (see Shadish, Cook, & Campbell, 2002), almost all reference the original work of Campbell and Stanley. Therefore, I chose to do the same in the present study. Often, throughout many fields of study, intervention research is conducted as a true experiment in which participants are randomly assigned to groups. This type of experimentation allows researchers to establish causal relationships among various phenomena. However, true experimental intervention research is nearly impossible to implement in typical school classrooms (de Anda, 2007).

In order to conduct a true experiment, the researcher or researchers must be able to fully control all variables involved in a study (Campbell & Stanley, 1963). A true experiment must allow for the random assignment of participants to groups and manipulation of the independent variable. Ideally, all experiments would be true experiments, allowing researchers to make strong causal claims about the relationships among various phenomena (de Anda, 2007). These true experiments are said to have strong internal validity, meaning that the differences in outcomes across groups are likely due to the intervention, and no other factor. Due to the nature of schools in which students are almost always pre-assigned to classrooms, thereby not allowing random
assignment of participants to groups, conducting a true educational experiment can be an impossible task.

Several decades ago, Campbell and Stanley (1963) suggested an alternative to the true experimental design called quasi-experimental design. Quasi-experiments occur when researchers are able to manipulate the independent variable, but are unable to randomly assign participants to groups as required for true experiments. While quasi-experiments are not as rigorous as a true experiments, and thus causal claims resulting from quasi-experiments are often subject to more scrutiny than those resulting from true experiments, a quasi-experimental research design does allow for manipulation of the independent variable and is, thus, superior to non-experiments when trying to make an argument for causality. One type of quasi-experimental design is a control group design. Campbell and Stanley (1963) stated:

One of the most widespread experimental designs in educational research involves an experimental group and a control group both given a pretest and a posttest, but in which the control group and the experimental group do not have pre-experimental sampling equivalence. Rather, the groups constitute naturally assembled collectives such as classrooms, as similar as availability permits but not so similar that one can dispense with the pretest (p. 47).

Conducting quasi-experimental intervention research, such as the control group design, allows educational researchers to make arguments for causal claims among various phenomena. However, such arguments are stronger when researchers have addressed potential threats to internal validity (Campbell & Stanley, 1963). Utilizing a pretest in a control group design allows for an examination of whether differences in posttest scores were due to differences in initial group make up, a threat to internal validity called selection. Some of the more common internal validity threats include:
1. Differential attrition of participants: This can lead to a large imbalance in the number of control and treatment group participants who complete the study, thus skewing outcome results. Another problem associated with participant attrition is that outcome differences between control and treatment groups may be due to fundamental differences among participants who remained in the study as opposed to the treatment itself (i.e., did those who completed the study have some characteristic that led them to remain in the study, thus affecting the outcome?). Repeated testing: When participants are tested several times over the course of a study, they may recall correct responses for future tests. Participants may also become conditioned to know that they are being tested. These issues can affect test scores and result in outcomes that are not necessarily due to the treatment.

2. Selection bias: Outcome differences between the treatment and control groups may be due to prior differences between groups. That is, before the beginning of the study, the groups may possess some characteristic that makes them incomparable. Thus, some factor other than the intervention itself may influence the posttest results for the two groups.

Pressley and colleagues (2006) conducted a meta-analysis of contemporary educational intervention research. Through their analysis, they determined several characteristics present in these empirical studies. They also made suggestions for the future path of educational intervention research. The authors argued that future educational intervention research must include complex treatments administered over several years, yet must remain ethical and methodologically sound in order to be received
as credible by the education community. They also encouraged educational intervention researchers to gather all of the data that is available, advising that using multiple data sources is often more revealing about the influence of an intervention. Finally, and in agreement with Hsieh et al. (2005), Pressley and colleagues (2006) attested that intervention research is highly under-utilized in the field of education. Pressley and colleagues contended that students in classrooms would benefit from context and content specific interventions. The authors encouraged educational researchers to conduct more intervention studies and to effectively communicate their results and suggestions for improvement to the educational community.

Quasi-experimental intervention studies are often quantitative in nature (Pressley et al., 2006). However, Pressley and colleagues posited that including a qualitative component within an intervention study may inspire new ideas and provide a better understanding of how and why the intervention(s) worked, beyond what can be understood from quantitative analysis alone. Thus, it seems that conducting mixed methods intervention research, such as the present study, can provide richer results than quantitative research alone.

The collection of additional qualitative data, as recommended by Pressley and colleagues’ (2006), can help to inform how interventions should be implemented and improved to ensure efficacy of the treatment. These qualitative data can be used to make immediate modifications and are also beneficial when considering changes to future implementations. This process is similar to the iterative intervention task improvement process of design research (Gravemeijer & van Eerde, 2009).
Conclusion

Much has been written about self-efficacy beliefs and collaborative learning. Yet insufficient research has been conducted that investigates the relationship that exists between the two. In order to help students increase their self-efficacy beliefs and, in turn, achieve at higher academic levels, educators should be presented with a clear understanding of how collaborative classroom learning can influence student self-efficacy beliefs. It has been shown that collaborative learning promotes higher academic achievement than competitive or individual learning (Johnson & Johnson, 1989). This may be due, in part, to the effect that high achieving group members’ academic success has on the self-efficacy beliefs of other group members (Bandura, 1986). Bandura (1995) believed that peer modeling, or vicarious experience, is one means of increasing student self-efficacy beliefs. By observing high achieving peer models in collaborative learning settings, students can increase their self-efficacy beliefs, thus promoting greater academic performance.

The results of contemporary empirical studies have demonstrated that statistics students, even at the college level, continue to struggle to understand basic statistical topics (Garfield, 1995). Undergraduates studying statistical sampling for the first time often find the topics associated with this field of study to be very complex and abstract (Nguyen, 2005; Yilmaz, 1996). For many students, this is the first and only time they will be presented with material regarding statistical sampling, making it difficult for instructors to help students understand the topics associated with this field of study (Fecso et al., 1996; Yilmaz, 1996). Several statistical sampling scholars and instructors have suggested that providing students with real world, hands-on tasks and activities can
help non-specialist undergraduate statistics students gain a clear, conceptual understanding of the various complex concepts associated with statistical sampling (see Chang et al., 1992; Fecso et al., 1996; Gelman & Nolan, 2002; Hodgson & Burke, 2000; Kelly, 2010; Mills, 2002; Richardson, 2003; Warton, 2007; Yilmaz, 1996).

In the present study, I hypothesized that peer modeling, a phenomenon that has been shown to influence students’ self-efficacy beliefs, is the mechanism through which collaborative learning influences the academic achievement of students learning about statistical sampling. That is, through this intervention study I attempted to establish self-efficacy beliefs as the mediating factor between collaborative learning and academic achievement. In order to test for the causality among collaborative learning, self-efficacy beliefs, and academic achievement in statistical sampling, I conducted an intervention study to pursue answers the following research questions:

1. Does working in small, collaborative, student led, peer groups in the classroom increase the topic specific self-efficacy beliefs of undergraduate level statistics students?

2. Do topic specific student self-efficacy beliefs mediate the relationship between collaborative learning and the academic achievement of undergraduate level statistics students studying statistical sampling?

3. How should the collaborative tasks intended to help undergraduate students learn topics related to statistical sampling be redesigned and implemented such that they foster the topic specific self-efficacy beliefs and, thereby, promote higher academic achievement of undergraduate students?
4. How should the overall research design be revised for use in future experiments of the same type implemented in similar contexts?

Because intervention studies are often difficult to conduct in an educational setting (de Anda, 2007), I used qualitative research methodology to assess and improve the collaborative learning tasks that comprised the intervention implemented with the students in the treatment group. Using this method of study also allowed me to provide a springboard for teachers and researchers who wish to implement a similar study in their classrooms. In the following chapter, I outline the research methodology I used to investigate the relationships among collaborative learning, self-efficacy beliefs, and academic achievement among college statistics students learning about statistical sampling.
CHAPTER 3
METHODS

The goal of this intervention study was to use quantitative and qualitative methods to examine the relationships among student self-efficacy beliefs, collaborative learning, and the academic achievement of students studying statistical sampling. More specifically, using quantitative methods, I attempted to determine if working in small collaborative groups in a college statistics classroom could help increase student self-efficacy beliefs for specific statistical concepts associated with statistical sampling, thus establishing self-efficacy as the mediating mechanism between collaborative learning and academic achievement for undergraduate statistics students. I employed qualitative research methods to study how the participants were experiencing the intervention, and to help ensure that the intervention was administered according to plan. Further, the data collected and analyzed using qualitative methods allowed me to revise the collaborative intervention tasks during the study, a method akin to the iterative processes used in design research (Cobb, McClain, & Gravemeijer, 2003; Gravemeijer & van Eerde, 2009; van den Akker, Gravemeijer, McKenney, & Nieveen, 2006). Using qualitative research methodology also provided additional insight into students’ understanding of and ways of learning about statistical sampling. Finally, qualitative methods helped to provide richer description of the intervention and its effects that may prove useful to teachers or researchers who choose to implement similar collaborative learning interventions in their classrooms. In this section, I describe the methodology I used to investigate the
relationships among collaborative learning, self-efficacy beliefs, and the academic achievement of students learning about statistical sampling.

**Participants**

During the 2011 Fall semester, I taught one instructional unit focusing on the topic of statistical sampling to 61 college students enrolled in two sections of a Basic Statistics course at a university located in western North Carolina. The students enrolled in these two sections of the course were not majoring in a mathematical or scientific field of study, and are primarily Political Science and Criminal Justice majors. According to the university registrar, the course was:

An introduction to statistical problem solving. Topics include organization and presentation of data; measures of location, variation, and association; the normal distribution, sampling distributions, and statistical inference. Emphasis will be on conceptual understanding and interpretation of results rather than theoretical development.

The 61 students enrolled in these two sections of the course were divided into two classes of 29 and 32 students. Because the students had already self-selected into each of the two course sections, I did not assign students to groups. That is, the students were separated based upon the section of this course in which they enrolled. Therefore, this was a quasi-experimental intervention study. One class served as the control group and the other as the treatment group for the study. I was not the primary instructor for these courses, however the same cooperating professor was the instructor of record for both sections of the course.

Of the 61 students enrolled between the two classes, 41 students (67% of total class) consented to participate in the study, 18 control class students and 23 students in the treatment class. All 41 students completed the pre unit survey and assessment and 27
of the participating students completed the post unit survey and assessment. That is, 13 students in the direct instruction class completed both pre and post surveys and assessments and 14 students in the collaborative learning class completed all pre and post paperwork. Therefore, only these 27 cases are included in the path model and mediation analysis, as well as the computations of descriptive statistics.

The gender of participating students was fairly evenly split between male (56%) and female (44%). The majority of students participating in the present study reported their race or ethnicity as white (95%), with Hispanic or Latino (2%) and African American (2%) being the only other races reported. Most of the students were of Junior class status (74%), along with a few Sophomores (15%) and Seniors (11%). This Basic Statistics class was designed primarily for non-mathematics and science majors, and the majority of participating students listed their major as Criminal Justice (56%). The other majors represented were Political Science (22%), Nursing (11%), Construction Management (4%), Public Administration (4%), and Communications (3%). The participating students were of typical undergraduate student age. The average age of participating students was 20.2 years old, with a minimum age of 19 and a maximum of 22 years old. Table 1 displays these demographic data separated by control and treatment classes:
Table 1

*Student Demographics*

<table>
<thead>
<tr>
<th>Control Class</th>
<th>Gender</th>
<th>Race*</th>
<th>Class</th>
<th>Major</th>
<th>Age*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>40% Male</td>
<td>100% White</td>
<td>17% Soph.</td>
<td>38% Criminal Just.</td>
<td>20.3 (0.75)</td>
</tr>
<tr>
<td></td>
<td>60% Female</td>
<td>57% Juniors 26% Seniors</td>
<td>24% Political Science 18% Nursing 7% Communications 3% Const. Mgmt. 3% Public Admin. 3% Athletic Training 3% Child Development</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Treatment Class</th>
<th>Gender</th>
<th>Race*</th>
<th>Class</th>
<th>Major</th>
<th>Age*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>58% Male</td>
<td>92% White</td>
<td>13% Soph.</td>
<td>55% Criminal Just.</td>
<td>20.1 (0.66)</td>
</tr>
<tr>
<td></td>
<td>42% Female</td>
<td>4% Af. Amer. 4% Hisp/Lat.</td>
<td>68% Juniors 19% Seniors</td>
<td>26% Political Science 6% Exercise Science 3% Nursing 3% Communications 3% Psychology 3% Risk Management</td>
<td></td>
</tr>
</tbody>
</table>

*Note. These averages pertain only to students who consented to participate in the study.*
Table 2 displays the breakdown of student attendance for each day of the study.

Table 2  
**Student Attendance by Class Period**

<table>
<thead>
<tr>
<th>Class</th>
<th>Attendance</th>
<th>Gender</th>
<th>Race</th>
<th>Class</th>
<th>Major</th>
<th>Age (M)</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27(93%)</td>
<td>48% F</td>
<td>97% W</td>
<td>74% Jr.</td>
<td>81% CJ/PS</td>
<td>20.8</td>
<td>13 of 13</td>
</tr>
<tr>
<td>2</td>
<td>24(83%)</td>
<td>54% F</td>
<td>100% W</td>
<td>88% Jr.</td>
<td>75% CJ/PS</td>
<td>20.3</td>
<td>12 of 13</td>
</tr>
<tr>
<td>3</td>
<td>22(76%)</td>
<td>50% F</td>
<td>95% W</td>
<td>72% Jr.</td>
<td>82% CJ/PS</td>
<td>20</td>
<td>13 of 13</td>
</tr>
<tr>
<td>4</td>
<td>25(86%)</td>
<td>60% F</td>
<td>92% W</td>
<td>64% Jr.</td>
<td>68% CJ/PS</td>
<td>20.2</td>
<td>13 of 13</td>
</tr>
<tr>
<td>5</td>
<td>21(72%)</td>
<td>52% F</td>
<td>100% W</td>
<td>86% Jr.</td>
<td>86% CJ/PS</td>
<td>20.9</td>
<td>12 of 13</td>
</tr>
<tr>
<td>6</td>
<td>18(62%)</td>
<td>50% F</td>
<td>94% W</td>
<td>83% Jr.</td>
<td>78% CJ/PS</td>
<td>20.3</td>
<td>13 of 13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class</th>
<th>Attendance</th>
<th>Gender</th>
<th>Race</th>
<th>Class</th>
<th>Major</th>
<th>Age (M)</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25(78%)</td>
<td>44% F</td>
<td>92% W</td>
<td>88% Jr.</td>
<td>76% CJ/PS</td>
<td>20.1</td>
<td>14 of 14</td>
</tr>
<tr>
<td>2</td>
<td>23(72%)</td>
<td>43% F</td>
<td>96% W</td>
<td>87% Jr.</td>
<td>91% CJ/PS</td>
<td>20.7</td>
<td>14 of 14</td>
</tr>
<tr>
<td>3</td>
<td>22(69%)</td>
<td>41% F</td>
<td>91% W</td>
<td>72% Jr.</td>
<td>82% CJ/PS</td>
<td>20.4</td>
<td>13 of 14</td>
</tr>
<tr>
<td>4</td>
<td>20(63%)</td>
<td>40% F</td>
<td>100% W</td>
<td>80% Jr.</td>
<td>75% CJ/PS</td>
<td>20.1</td>
<td>14 of 14</td>
</tr>
<tr>
<td>5</td>
<td>21(66%)</td>
<td>48% F</td>
<td>95% W</td>
<td>76% Jr.</td>
<td>90% CJ/PS</td>
<td>20.1</td>
<td>13 of 14</td>
</tr>
<tr>
<td>6</td>
<td>20(63%)</td>
<td>40% F</td>
<td>90% W</td>
<td>70% Jr.</td>
<td>75% CJ/PS</td>
<td>20.5</td>
<td>14 of 14</td>
</tr>
</tbody>
</table>

*Note. F = Female, W = White, Jr. = Junior, CJ = Criminal Justice, PS = Political Science. ‘Attendance’ refers to the number of total students attending class, regardless of participation status. ‘Participants’ refers to the number of participants attending each day.*

Exposure to the intervention was rather high for participants in the treatment class. In the treatment group, 85.7% of participants attended all six class periods, with a range of 13 to 14 participants attending each day. Similarly, in the control group, 84.6% of participants attended all class periods, with a range of 12 to 13 participants per class.

In order to encourage participation in this study and to establish familiarity with the students, I attended several class meetings over the three weeks prior to the day of recruitment and beginning of data collection. The intention was to help the students feel more comfortable with my presence in the classroom. I believed that my regular
attendance in the classroom would help to establish familiarity and would encourage a higher level of participation on the part of the students.

On the day of recruitment, the first day of data collection, a total of 52 students attended the two classes. In the control group, 27 students attended the first day of class and 25 students in the treatment class were in attendance. During the recruitment period and all five subsequent class periods, the cooperating professor was not present in the classroom. I worked from a recruitment script (see Appendix A) to explain why I was there and, briefly, the intentions of the study. I tried to be very clear that participation in the study was both voluntary and confidential (i.e., the cooperating professor would never know who did or did not participate in the study). I also stated that there were no rewards for participation and no penalties for choosing not to participate, and that students who initially chose to participate could opt out of the study at any time. Then, I gave each student, regardless of her or his willingness to participate, a packet of materials containing a description of the study and consent form (see Appendix B), a self-efficacy beliefs survey (see Appendix C), and a prior statistical knowledge assessment (see Appendix D). Each packet was placed in a brown folder. Giving every student a materials packet allowed for discretion regarding participation. The students who chose to participate were asked to complete all three forms. Those who chose not to participate were asked to complete the consent form, noting on the form that they did not wish to participate in the study. Once all students had completed the appropriate paperwork, the packets were collected. Further detail about the procedures of this study, including the description and use of the specific measures, are described later in this chapter.
Classroom Context

Classroom Layout

The classroom in which the intervention took place was designed in a small auditorium-style layout. There were three rows of fixed, immovable tables, each placed on a riser such that every row was higher than the row in front of it. The rows were slightly curved in arc-like manner. The students’ seats were bolted to the tables and to the floor, and were therefore also stationary. This rigid architectural structure impacted the collaborative learning environment in various ways that are discussed in later chapters.

Establishing Rapport with Students

As noted, I was not the primary instructor for the two Basic Statistics courses in which the participants of this study were enrolled. Because this was the case, I took measures that I thought would help the students in both classes feel comfortable participating in this intervention study. In order to create a sense of familiarity and comfort, I attended both classes for three weeks prior to intervention implementation. However, I believe that the students in both the treatment and control groups struggled to become comfortable with and invested in the present intervention study because I was not the primary instructor for the classes, an issue that is discussed in later chapters.

Measures

Self-efficacy

For this study, I used a self-efficacy measure based on the Motivated Strategies for Learning Questionnaire (MSLQ) designed by Pintrich, Smith, Garcia, and McKeachie (1991) (see Appendices C and E). This questionnaire has been used as the survey
instrument for numerous studies investigating self-efficacy beliefs and there is evidence that inferences from scores from this measure have sufficient validity and reliability when used with samples similar to the one in the present study (Pintrich, Smith, Garcia, & McKeachie, 1993). The original MSLQ consists of 81 questions that cover various motivational constructs and strategies for learning. I chose to use seven of the eight Self-efficacy for Learning and Performance questions from the survey and amended these questions to be more specific to the instructional unit on statistical sampling. According to Duncan and McKeachie (2005), the coefficient of reliability, α, for the self-efficacy section of the MSLQ is 0.93. However, studies conducted in which only college students completed the self-efficacy portion of the MSLQ reported this value as ranging from .73 to .91 (Burlison, Murphy, & Dwyer, 2009; Duijnhouwer, Prins, & Stokking, 2010; Jacobson & Harris, 2008; Radovan, 2011). I expected the pre and post self-efficacy coefficient of reliability values to be within this range for the present study and report the coefficient of reliability in the Results section below. A similar self-efficacy beliefs questionnaire was administered both pre-instruction (see Appendix C) and post-instruction (see Appendix E), with the wording changed to reflect that one was designed to tap into student beliefs before the participants began the instructional unit and the other focused on student beliefs after instruction had occurred. The wording changes were very minor, simply changing the verb tenses from future to present for most of the items. For example, item three changed from “I’m confident I can understand the basic concepts regarding statistical sampling” on the pre-instruction survey to “I’m confident I understand the basic concepts taught regarding statistical sampling” on the post-instruction survey.
Prior statistical knowledge

Before providing any instruction or collaborative learning activities, the participating students in each class completed a pre statistical knowledge assessment (see Appendix D). This assessment focused solely on the students’ knowledge of statistical sampling. I developed this assessment based upon the required text for this course, ensuring that I used the same language as the author of the textbook. Also, the cooperating professor reviewed the assessment prior to the recruitment of students. Because I developed the prior statistical knowledge measure and, hence, there was no prior associated reliability data, I determined the reliability of the measure after the intervention was complete and all data had been collected. This value is reported in the Results section below.

Post unit statistical knowledge and understanding

After the students and I had completed the instructional unit on statistical sampling, the statistical knowledge and understanding of the participating students was assessed using a post statistical knowledge assessment (see Appendix D) identical to the pre statistical knowledge assessment. Because I developed the post unit statistical knowledge measure and, hence, there was no prior associated reliability data, I determined the reliability of the measure after the intervention was complete and all data had been collected. Findings regarding reliability, and their implications for interpretation of effects, are discussed in detail in the Results section.

Procedure

During the 2011 Fall semester, I taught the same instructional unit regarding statistical sampling to two sections of an undergraduate Basic Statistics course at a
university in western North Carolina. I served as both the instructor and sole researcher for the present study. The instructional unit was taught over six 50-minute class periods. The two classes met from 11am to 11:50am and 12pm to 12:50pm on Monday and Wednesday of each week, for a period of three weeks. The students also met with the cooperating professor for 50 minutes on Friday of each week in a computer lab environment. The Friday class meetings were intended to give students an opportunity to apply what they had learned in a more real world, hands-on setting. I was not present at the Friday class meetings. During the time between recruitment of participants and administration of the post statistical knowledge assessment, the cooperating teacher was not present at the Monday and Wednesday class meetings. This was to protect the confidentiality of the participating students. I used a random number generator to determine which class would serve as the treatment group and receive the collaborative learning intervention. The 11am class served as the control group and the 12pm class served as the treatment group for this study.

Prior to any instruction or collaborative group work, participating students completed a self-efficacy beliefs survey (see Appendix C) and a pre statistical knowledge assessment (see Appendix D). Once all participating students had completed the survey and assessment, I began instruction on the statistical sampling unit in the control class and started with a discussion of collaboration in the treatment class.

**Control Class Format**

In the control class, I taught using a traditional, lecture-style format (i.e., direct instruction). I encouraged students to ask questions and to participate in whole class discussion, but no collaborative group activities were administered to the students in this
These students did, however, complete individual activities periodically throughout the instructional unit. These activities were based upon assignments from the required text. There was no discussion of collaborative learning in the control class.

**Collaborative Learning Treatment**

**Student groups.** In order to help foster collaboration among the students in the treatment class, I tried to employ many of the recommendations provided by Cohen (1994b) in her book, *Designing Groupwork*, including how to group students, taking the architecture and setup of the classroom into consideration, and teaching students about collaboration and how to collaborate effectively. Originally, I had planned to assign students to groups of four. However, after realizing that the architectural setup of the classroom would not allow for effective collaboration among four students, I amended this plan and decided to pair students in dyads. This arrangement presented many issues that are discussed in later chapters. Having observed the students for several class periods prior to the first day of the instructional unit, and based upon discussions with the cooperating professor, I had a sense of the academic performance level and social interaction skills of each student in the treatment class. I also had a list of students’ class (i.e., Freshman, Sophomore, Junior, or Senior) and major. Therefore, I was able to pair the students in a manner that I believed would help foster collaboration and learning. The initial criteria I used to pair students was major, because I believed the students would benefit from completing some of the upcoming tasks with peers within the same major course of study. Next, per Cohen’s (1994b) recommendation, I considered each student’s level of academic performance in this class. Various scholars (see Cohen, 1994b; Johnson & Johnson, 1989) have recommended pairing middle performing students with
high performing students, and middle performing students with low performing students. These scholars have shown that this type of pairing can help alleviate frustrations high or low performing students may experience when working with a peer who is performing at an academic level vastly different from their own.

**Foundational learning theory.** Because this intervention study was informed by the ideas of design research (Cobb et al., 2003), I chose to establish a learning theory to use as the foundation from which to develop the daily collaborative tasks and class structure. Based upon the GAISE College Report (2005), suggestions from various scholars in the field of statistics education (i.e., Ben-Zvi, 2005; Ben-Zvi & Garfield, 2005; delMas et al., 2007; Garfield, 1995; Garfield, 2005), and my observations of the class as the students were covering the material prior to implementation of the collaborative learning intervention, I believed that the students in the treatment class would begin this instructional unit with a basic understanding of a sample and a population, but would have minimal prior knowledge applicable to this unit regarding statistical sampling. I also believed, based upon the literature reviewed above, that these students would have a certain level of anxiety and fear of this statistics class and of the topics covered in this instructional unit on statistical sampling. As several statistics educators have noted, most introductory statistics students simply want to find the correct answer and pass the assignment or exam (Ben-Zvi & Garfield, 2005; Garfield, 1995; Garfield & Ahlgren, 1988). Based upon the results of contemporary empirical literature regarding statistics education and the suggestions made by statistics educators, I believed the students’ learning trajectory would take the following path:
1. Students would begin this unit on statistical sampling with a very simple understanding of sample and population, with little prior knowledge of the other concepts to be covered in this unit.

2. Students would feel anxious about their performance on the tasks and assignments associated with this unit and their ability to understand the topics in this instructional unit.

3. During the first few days of the instructional unit on statistical sampling, students would be more focused on finding the correct answer than on acquiring statistical understanding and reasoning skills.

4. Through the use of collaborative group work in the classroom, during which time students were given the opportunity to discuss their ideas with a small group of their peers, students would begin to develop a sense of statistical understanding and would gain the ability to reasonably discuss statistical ideas.

In order to overcome students’ quest for the right answer, The GAISE College Report (2005) suggested that teachers should focus on fostering statistical understanding and reasoning in the classroom. Based upon the suggestions from statistics education experts and the literature regarding collaborative learning, the foundational learning theory included the following:

1. With the treatment group of students, I would begin by discussing what the term collaboration meant, how collaborative groups could be effective and productive, and what was expected from each member of a collaborative group. I did not expect that all students would have the skills necessary to engage in collaborative
learning and, thus, followed the advice of Cohen (1994b) to discuss these expectations and skills prior to the collaborative learning intervention.

2. Daily, I would use collaborative learning as an instructional method to allow students to work, discuss, and collaborate with their peers. I believed this type of learning environment would give students the freedom to express their thoughts and ideas, feel safe to be wrong, and generate their own comprehension of statistical topics. Allowing for these periods of discussion, negotiation, and argumentation should promote collaborative learning within the student groups (Chinn, 2006; Cohen, 1994b; Dillenbourg, 1999; Erkens et al., 2006; Grover, 2010; Johnson & Johnson, 1989; Ploetzner et al., 1999). I also believed that this opportunity to collaborate with their peers would help students feel a greater sense of self-efficacy regarding their statistical knowledge and understanding of the concepts associated with statistical sampling, as well as their ability to reason about statistical ideas. This collaborative learning time would serve as the primary means of instruction for the students, with little to no intervention from me, the instructor.

3. I would allow a short period of whole class discussion at the end of each class. The intent of this debriefing period as recommended by Hodgson and Burke (2000) was to ensure that the students understood the material as I expected, to share insights the students had gained while working collaboratively with their peers, and to answer any outstanding questions regarding the statistical topic of the day.
Iterative cycles. The iterative process of task implementation, observation, reflection, implementation, etc. used in the present study was informed by the theoretical and empirical design research literature (Cobb et al., 2003; Gravemeijer & van Eerde, 2009; van den Akker et al., 2006). The qualitative portion of the present study, while not a true design experiment, was influenced by and had several similarities to the processes of design research. The iterative cycles of classroom observation and subsequent reflection helped to highlight areas in which the daily collaborative tasks should be modified and served to improve the efficacy of the collaborative intervention treatment as it was implemented in this study. By dynamically studying and improving the collaborative tasks, I tried to ensure that the students in the treatment class were truly experiencing collaboration. During the final retrospective analysis, I was able to identify improvements that should be made to the overall design and bring to the fore recommendations for future research studies of this type. These changes are discussed in subsequent chapters.

Daily class periods and task implementation. On the first day of the instructional unit, after all students in the treatment class had completed the consent paperwork and pre survey and assessment, I began class with a discussion of what collaboration is and what it is not. I let the students lead this discussion, prompting them with a series of questions regarding their historical experiences with this type of learning. Most students understood that collaboration meant working with others. When pushed to distinguish collaboration from cooperation, only one student came up with the idea that collaboration includes creating common ideas, thoughts, products, etc. I emphasized that collaborative learning means that pairs or groups of students work together to create
mutually agreed upon products, and that these products can take the form of ideas, definitions, portfolios, performances, completed worksheets, etc. Once it was clear that most of the students in the treatment class were comfortable with our common definition of collaboration, I had the students engage in a collaborative task unrelated to the instructional unit on statistical sampling.

The students, grouped in pairs as mentioned previously, were asked to compose a list of at least three questions they would ask if they had the opportunity to interview someone they admired. During my initial observations, the students were inclined to create three questions each, uniquely pertaining directly to the person they would interview. However, I interjected and asked each pair of students to come up with a list of questions that would be suitable for either admirable person being interviewed. Therefore, each pair of students had to discuss who they would like to interview, what they would like to know about that person, and how they could come up with one common list of questions that could potentially elicit the answers they sought from either admirable person. This required students to think about the intention of their questions and pushed them to develop questions worded in a way that would serve both students. This collaborative learning task lasted for approximately 12 minutes. Once all pairs of students had completed this task, the full class and I once again discussed the idea of collaboration. After answering the few remaining questions regarding collaborative learning, we moved onto the first tasks designed for this instructional unit on statistical sampling.

Each of the tasks developed for this intervention study were designed to follow a similar implementation format (see Appendices F through J). The implementation
sequence contains three phases and is called Launch, Explore, Summarize. This instructional model was developed as a method of implementing problem-based curricula (Connected Mathematics Project, 1995). During the Launch phase, the instructor explains the task and sets appropriate expectations for the students. The instructor must be careful not to provide students with too much information so that students are appropriately challenged by the task. The Explore phase is the period when students work, usually in small, collaborative groups, to complete the task. During this phase, students discuss the problem with their partner(s), offer ideas and solutions, and negotiate with one another to come to an agreed upon final outcome. The final phase, the Summarize phase, is a time for the students and instructor to come together as a group and engage in a final debriefing session. During this time, students discuss their experiences, present their ideas, offer alternative solutions, and ask any lingering questions. This is also a time when the instructor formatively assesses student understanding and knowledge and offers any necessary clarifications. The Launch, Explore, Summarize process for each of the collaborative learning tasks implemented during the present study is described in full detail in Appendices F through J.

In the present study, the Launch phase of some classes began with a brief introductory whole class discussion regarding the topic of the day. However, during the Launch phase of other class periods, I asked the students to engage with their partner and begin work on the daily tasks prior to any introductory whole class discussion. My intention with this alternative method was to determine how the students would interact with one another, having little to no direction from or engagement with me, the
instructor. I wanted to understand which class format seemed to work better for this
group of students.

Regardless of whether or not we engaged in introductory whole class discussion
during the Launch phase, once the students were ready to begin work on the daily tasks, I
presented the collaborative learning activity for the day. If a handout or worksheet was
available to guide students through the tasks, I gave each dyad only one copy of the
handout or worksheet. According to Cohen (1994b), this is one way to encourage
students to work collaboratively.

During the Explore phase, I moved throughout the classroom observing pairs of
students as they worked together and listening for any glaring misconceptions the
students seemed to have. Similar to the process of design research (Gravemeijer & van
Eerde, 2009), I took copious notes, recording any positive responses, problems, issues,
suggestions, questions, etc. presented by the students during this phase. I also noted any
problems or potential obstacles concerning task design, especially focusing on the
influence tasks appeared to have on students’ understanding of statistical sampling. In
order to assess whether or not collaboration was occurring, I specifically looked for the
criteria noted by various scholars in the field of collaborative learning (see Chinn, 2006;
Cohen, 1994b; Dillenbourg, 1999; Erkens et al., 2006; Grover, 2010; Johnson & Johnson,
1989; Ploetzner et al., 1999). I observed to see if students were engaging in thoughtful
discourse and using appropriate academic language while arguing their points and
negotiating with their partner. I noted whether or not students were helpful to their
partner, working together to reach the task goal(s), thus fostering an atmosphere of
positive social interdependence. I also tried to determine whether or not students seemed
comfortable interacting with their peers in this manner. Using the data from these observations, I was able to appropriately revise the tasks and class format for subsequent class periods.

Following the suggestion of various scholars in the field of cooperative and collaborative learning (Cohen, 1994b; Johnson & Johnson, 1989), I only intervened with the students during the Explore phase if I believed that they were far off topic, either due to misunderstandings or due to a loss of focus. At times, I brought the students back together for whole class discussion or large group collaboration in order to help them complete the tasks successfully. Both of these alternative methods of instruction are presented and discussed in the Results and Discussion chapters. During the final Summarize phase of each class period, the students and I engaged in whole group discussion to address any remaining questions or problems and to allow students to share their insights or contributions regarding the topic of the day.

The structure described was the general flow for each class period. However, I made changes to this general structure based upon my daily observations during each class period and the subsequent reflective periods. This iterative process allowed me to ensure that I was offering students the best collaborative learning experience possible. In the Results chapter, I offer further detail of each class period and information regarding diversions from the basic class structure described above.

**Task development.** Based upon the recommendations of several college statistics instructors and scholars, (see Chang et al., 1992; Fecso et al., 1996; Gelman & Nolan, 2002; Hodgson & Burke, 2000; Kelly, 2010; Mills, 2002; Richardson, 2003; Warton, 2007; Yilmaz, 1996), hands-on, real world activities were developed to help
non-specialist students gain a conceptual understanding of the major ideas associated with statistical sampling. Current experts in the field of statistics education were consulted and provided feedback to help strengthen the efficacy of the intervention tasks. Below, each concept addressed in this instructional unit is briefly described and the means by which these concepts were developed through instruction are outlined. The general format for implementing the tasks is discussed in further detail in Appendices F through J.

**Statistical sampling concepts.** In this instructional unit, the following concepts related to statistical sampling were addressed:

1. **What is a sample?**
   
   A sample is a subset of the population. A sample is studied in order to make accurate inferences about the population.

2. **What is a population?**
   
   A population is the entire collection of units (e.g., people, animals, M&Ms, etc.) from which data can be collected. If data are gathered from the entire population, this is called a census.

3. **What distinguishes a sample from a population?**
   
   A sample is only a portion of a population and is, therefore, smaller than the population that is being studied. In most studies, data are gathered from a sample and are analyzed to make inferences about the population. Because the entire population data are not typically analyzed, sampling bias is a factor when inferring from a sample to a population. Sampling bias occurs because a sample has been chosen to represent the population. This type of bias is expected and is
accounted for with the margin of error, or the estimated difference between the sample statistic and the population parameter.

4. How are accurate inferences made from a sample to a population?

In order to make accurate inferences about the population, a sample must reflect the characteristics of the population as closely as possible. The sample must also be of appropriate size in order to ensure that accurate inferences can be drawn from the sample to the population. To choose an appropriate sample, researchers must understand which characteristics are important to the study and must devise a sampling method that will create a sample that very closely resembles the population with regard to those characteristics. Once a sample is chosen, there are various analyses that can be computed from the sample data to make inferences about the population. One of these is the computation of a confidence interval. A confidence interval includes a sample statistic (e.g., sample mean, sample proportion, sample standard deviation, etc.) and the margin of error, or the estimated difference between the sample statistic and the population parameter.

5. What are the different methods for choosing a representative sample?

There are several methods that can be used when choosing a sample that best represents the population. They are:

   a. Simple random sample: This method involves selecting a sample by randomly choosing each member of the sample such that each member of the population has an equal chance of being chosen for the sample. Also, each sample of size \( n \) has the same chance of being chosen. This can be done in a number of ways. The names or numbers of each member of the
population can be put in a hat and individual names or numbers can be
drawn at random until the number of members needed for the sample is
reached. Alternatively, a random number generator or table of random
digits can be used to choose a simple random sample.
b. Stratified random sample: Sometimes there are factors that divide the
population into different groups, or strata, upon which it is expected the
characteristic(s) of interest will vary. To ensure that the sample accurately
reflects the population, researchers have to account for when choosing a
sample. Stratified random sampling is a method in which the population
is first divided into strata, or sub-groups. Then, a simple random sample is
chosen from each of the different strata.
c. Cluster random sample: Similar to a stratified sample, the cluster
sampling method begins by dividing the population into groups, or
clusters. Then, a simple random sample of full clusters is chosen and all
members of the chosen clusters comprise the sample. Cluster random
sampling is typically used when researchers cannot get a complete list of
the members of a population they wish to study but can get a complete list
of groups, or clusters, within the population. This sampling method may
also be used when a simple random sample would produce a list of
members so widely scattered that surveying them would prove to be too
expensive and/or time consuming.
d. Systematic random sample: In systematic random sampling, the sample is
chosen from an ordered list of the members of a population, also known as
the sampling frame. In order to ensure that each member of the population has the same chance of being chosen, the researcher chooses a random starting point in the sampling frame and members of the population are chosen at some pre-determined sampling interval, \( k \). That is, each \( k^{th} \) member listed in the sampling frame is chosen as a member of the sample. \( k \) could indicate fifth, seventh, twenty-third, etc. Researchers often choose this method of sampling because of the ease in which a sample can be chosen and because of the periodic system involved.

e. Convenience sample: A convenience sample is not random and every member of the population does not have the same chance of being chosen to become part of the sample. However, this method of sampling is often used in political exit polling or to obtain a quick and easy survey of public opinion regarding a particular topic. In convenience sampling, a surveyor waits for people to pass by and stops them to ask questions of interest. While convenient, inexpensive, and easy, it is not recommended that researchers use convenience sampling if they want to create a sample that accurately reflects the population.

Single stage sampling includes only one of the methods listed above in a through e. Multi-stage sampling, however, combines two or more of the methods listed above. When it is necessary to comprise a very specific sample, often a multi-stage sampling method will be used.
6. Why does the sampling method matter?

In order to make accurate inferences about the population, a sample that best reflects the characteristics of the population must be chosen. If a sample is chosen that does not reflect the population, the inferences made based upon the sample data can be misleading.

7. Issues that can arise when sampling a population

When sampling a population, several issues can arise that may lead to invalid data and inaccurate inferences about the population. Researchers have to be careful not to introduce too much bias into the data collected. Sampling bias will occur simply because a sample, a smaller subset, has been chosen to represent the total population. This type of bias is expected and is accounted for with the margin of error, or the estimated difference between the sample statistic and the population parameter. However, other types of bias such as selection bias, response bias, and nonresponse bias can adversely affect the outcome of a study. Selection bias occurs when the sample is comprised of only people who have volunteered to participate in the study. It is possible that the people who volunteered to participate have a vested interest in the study or that they vary in some characteristic way from those who chose not to participate. Response bias occurs when a representative sample is chosen from the population, but sample participants provide incorrect information. Participants may not be comfortable with certain survey questions, they may be embarrassed to offer the truth, or they may simply lie when responding to certain survey items. Nonresponse bias occurs when a representative sample is chosen from the population, but members
of the sample cannot be contacted or choose to not respond. With the exception of sampling bias, or the bias that is expected when sampling a population, all types of bias can produce sample data that are not representative of the population.

8. Using sample surveys versus a census to estimate a population

It can be expensive and time consuming to conduct a census, or to survey the entire population of interest. Therefore, researchers commonly select a sample from which to gather data. These data are then used to make inferences about a population. It is understood that computed sample statistics are not exactly the same as population parameters. However, it is acceptable to use sample data to make inferences about a population, as long as the sampling methods are sound and produce a sample that accurately reflects the population. It is accepted that sampling bias will occur as a result of choosing a sample. The margin of error is accepted as an accurate estimate of the difference between a sample statistic and a population parameter. Overall, it is common practice to use sample data instead of a census to make inferences about the population.

**Concept development.** The content for the three-week instructional unit focused on topics related to statistical sampling. The unit included the above concepts. The intention of teaching these concepts within the same unit was to help students understand the overall concept of statistical sampling and to encourage them to think about gathering information about the whole (i.e., the population) by analyzing a subset of the whole (i.e., a sample).
The first concepts taught in this instructional unit were the difference between a sample and a population and the idea that a sample is a small portion of the entire population. While learning these concepts, students have the opportunity to consider how to gather a small portion of the population, or how to choose a sample, and then how to collect data from that sample. That is, students can determine if they would conduct a census, which can be time consuming and costly, or if it is more beneficial to choose a sample that best represents the population. If they decide to choose a sample, students need to determine how to choose a sample that best represents the population and how they will gather data from that sample. Will they choose a simple random sample, or will they use some other method of sampling such as stratified sampling, cluster sampling, systematic sampling, or convenience sampling? Students may decide to use the methods of stratified or cluster sampling if resources, such as time and money, are insufficient to survey a large number of participants. Choosing a representative sample is a key component of sound statistical sampling.

Once students determine how they choose a sample and collect data, they are ready to analyze the data. In this unit, students have the opportunity to work with both descriptive and inferential statistics and learn how to infer from sample data to the greater population. Students also have a chance to explore the data and to determine which statistics are descriptive, or those that describe the data set, and which statistics are inferential, or those from which they can infer something about the population. It is important that students understand the difference between these two types of statistics so that they do not use descriptive statistics to try to make statements about the population. By making inferences about the population and understanding this relationship between a
sample and a population, students begin to think about the concept of sampling bias, or the accepted difference between sample values (i.e., statistics) and population values (i.e., parameters), and why sampling bias exists.

This understanding of sampling bias prompts students to begin thinking about the concept of margin of error, or how well the sample allows one to make inferences about the population. In the text for this course, the formula for the conservative margin of error, $\frac{1}{\sqrt{n}}$, is used. Exploring the concept of margin of error reinforces a sound understanding of the concepts of sample versus population, choosing a sample that best represents the population, and making inferences about the population from sample data. This is an appropriate time in the instructional unit to discuss current events and recent poll data students may have read or heard about. These discussions often lead to an understanding of basic confidence intervals, how they relate the sample to the population, and how they are relevant in current media. It is important to ensure that students understand how the concepts of sample, population, inferential statistics, sampling error, margin of error, and confidence intervals are related and how they relate to the overall topic of statistical sampling.

Because the margin of error is related to the sample size, this is a suitable time to introduce the concept of choosing an appropriate sized sample and reinforcing the need to choose a sample instead of gathering census data from the entire population. It is important that students understand that the larger the sample, the lower the margin of error, while also considering that choosing a large sample requires more resources (i.e., money, time, etc.). This relationship between sample size and margin of error can easily
be demonstrated by asking students to calculate the necessary sample size or margin of error for various scenarios and comparing the values.

The next two concepts, developing single stage and multi-stage sampling plans, help students connect all of the concepts they learned during the instructional unit thus far. By asking students to develop sampling plans, they are prompted to think about the population they wanted to sample, the kinds of data they wanted to gather from the sample and inferences they would like to make about the population, choosing an appropriate sized sample from which to collect data, and the associated sampling bias and margin of error. It is important that students develop a sound understanding of how to develop sampling plans and think about all of the previously mentioned criteria that should be considered. It is also critical that students recognize that gathering data from a sample that does not represent the population well can be the demise of a research study.

Once students have a good sense of how to develop sampling plans to gather data that are representative of the population, they begin to explore the process of developing survey questions and the issues that can arise during this process. For example, it is essential that students understand the difference between open questions that allow respondents to answer freely and closed questions that force participants to choose from a set of pre-scripted answers. It is also important that students understand when it is appropriate to use each type of question. Students need to understand that the way in which questions are worded will impact participants’ responses. Students then begin to explore the issues that arise when trying to measure attitudes and emotions, such as participants’ unwillingness to answer questions about certain controversial issues. In addition, certain questions may be embarrassing and could prompt participants to answer
in a way that is different from how they truly feel. Considering these issues that may arise when developing survey items can lead to discussions about response bias and nonresponse bias and the difference between the two. These discussions help students understand that response bias is the result of participants responding differently than they truly think or feel. Students also begin to understand that nonresponse bias is a form of bias that occurs when respondents differ in meaningful ways from those who did not respond for some reason (e.g., they chose not to respond, they could not be contacted, etc.).

This instructional unit was intended to help students develop a sound understanding of the concepts related to statistical sampling. For the treatment class, I developed collaborative tasks that I believed would help students understand these concepts and also help them understand the relationships among these concepts (see Appendices F through J for full lesson plans). In the control class, I asked students to work independently in class to complete activities and problems from the Utts and Heckard (2006) text, *Mind on Statistics*.

**Reflection and retrospective analysis.** Immediately following each class period I engaged in a period of reflection. I used this time to summarize the experiences of the day and how these experiences did or did not foster an atmosphere of collaborative learning. During this reflective period, I also determined how to appropriately modify the collaborative tasks and or overall class structure for the following class period(s) to help improve the intervention and ensure that collaboration learning was occurring. Once all six class periods were complete, I retrospectively analyzed all of the data gathered and noted any changes that served to improve the design of the experiment, the collaborative
learning tasks, and the foundational learning theory that served as the basis for this experiment. I hope to continue this line of research by conducting subsequent, similar research studies with undergraduate statistics students.

**Post Unit Survey and Assessment**

At the end of the instructional unit, the participating students in both classes completed the post self-efficacy beliefs survey (see Appendix E) and the post statistical knowledge assessment (see Appendix D). I distributed the survey and assessment, placed together in a brown folder, to all students regardless of participation status. This was intended to allow for discretion among classmates. Then, I collected the surveys and assessments from all students.

**Path Model**

Increased student self-efficacy beliefs can advance academic achievement (Bandura, 1977, 1995, 1997; Parajes, 1996; Parajes & Miller, 1997; Pajares & Urdan, 2006). Collaborative learning is also known to foster deeper intellectual understanding and to promote greater academic achievement (Cohen, 1994b; Johnson & Johnson, 1989, 2002, 2009). Based on the empirical evidence regarding the relationship academic achievement shares with both student self-efficacy beliefs and collaborative learning, I hypothesized and tested the path model shown in Figure 1:
Figure 1. Hypothesized path model.

The ‘collaborative learning’ variable shown on the left side of the model served as the treatment/control variable and is therefore a binary categorical variable.

Prior to the beginning of the intervention, using Mplus version 6.1, I conducted an a priori power analysis for the model shown above. I ran this analysis as a MIMIC model. Because I expected that a maximum of 30 students in the control classroom and 30 students in the treatment classroom would participate in the present study, I set the sample size at 60. I varied the effect size to reflect a small effect of 0.1, a moderate effect of 0.3, and a large effect of 0.5. I also varied the seed for each effect size to ensure that results were consistent (Muthen & Muthen, 2002). The results of the a priori power analysis are shown in Table 3:
Table 3
A Priori Power Analysis

<table>
<thead>
<tr>
<th>Relationship</th>
<th>0.1</th>
<th>0.3</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post Unit Knowledge and Pre Unit Knowledge</td>
<td>0.141</td>
<td>0.605</td>
<td>0.908</td>
</tr>
<tr>
<td>Post Unit Knowledge and Pre Self-efficacy</td>
<td>0.124</td>
<td>0.611</td>
<td>0.904</td>
</tr>
<tr>
<td>Post Unit Knowledge and Collaborative Learning</td>
<td>0.073</td>
<td>0.240</td>
<td>0.492</td>
</tr>
<tr>
<td>Post Unit Knowledge and Post Unit Self-efficacy</td>
<td>0.146</td>
<td>0.657</td>
<td>0.969</td>
</tr>
<tr>
<td>Post Unit Self-efficacy and Pre Unit Knowledge</td>
<td>0.123</td>
<td>0.649</td>
<td>0.966</td>
</tr>
<tr>
<td>Post Unit Self-efficacy and Pre Self-efficacy</td>
<td>0.136</td>
<td>0.622</td>
<td>0.970</td>
</tr>
<tr>
<td>Post Unit Self-efficacy and Collaborative Learning</td>
<td>0.077</td>
<td>0.225</td>
<td>0.501</td>
</tr>
</tbody>
</table>

As noted in Table 3, reaching an acceptable level of power would likely be an issue when a small or moderate effect was present. I recognized that, at these effect sizes, establishing statistical significance might prove difficult. I was unable to increase the size of the sample due to the enrollment limitation of 30 students per class. However, the qualitative research approach of the present study allowed me to make practical suggestions to improve the design of the experiment, the associated collaborative learning tasks, and to strengthen the foundational learning theory such that the resulting design and theory can be employed by researchers and practitioners in future experiments. I address the issues of establishing statistical significance and of statistical power in the Discussion chapter.

In order to test the hypothesis that self-efficacy mediated the relationship between collaborative learning and academic achievement for these students studying statistical
sampling, I used Mplus and its ability to conduct a Sobel (1982) test of mediation. Using Mplus to conduct this test produced an estimate of the indirect path from Collaborative Learning to Post Unit Statistics Knowledge mediated by Post Unit Self-efficacy Beliefs. It also provided a statistical test of the path estimate to determine statistical significance.

**Research Questions and Anticipated Results**

The purpose of this study was to attempt to answer the following research questions:

1. Does working in small, collaborative, student led, peer groups in the classroom increase the topic specific self-efficacy beliefs of undergraduate level statistics students?

2. Do topic specific student self-efficacy beliefs mediate the relationship between collaborative learning and the academic achievement of undergraduate level statistics students studying statistical sampling?

3. How should the collaborative tasks intended to help undergraduate students learn topics related to statistical sampling be redesigned and implemented such that they foster the topic specific self-efficacy beliefs and, thereby, promote higher academic achievement of undergraduate students?

4. How should the overall research design be revised for use in future experiments of the same type implemented in similar contexts?

The qualitative component of this study allowed me to highlight the areas of the research design that required revision. These included the collaborative learning tasks, the overall structure and format of the class, and the foundational learning theory. I was able to make practical suggestions to strengthen the learning theory regarding students’
knowledge and understanding of statistical sampling. College statistics instructors can use the revised collaborative learning tasks, modified learning theory, and overall study design as a foundation from which to develop and implement similar research experiments in their own classrooms.
CHAPTER 4

RESULTS

Through data collection and analyzing the results of the present study, I aimed to answer the following research questions:

1. Does working in small, collaborative, student led, peer groups in the classroom increase the topic specific self-efficacy beliefs of undergraduate level statistics students?

2. Do topic specific student self-efficacy beliefs mediate the relationship between collaborative learning and the academic achievement of undergraduate level statistics students studying statistical sampling?

3. How should the collaborative tasks intended to help undergraduate students learn topics related to statistical sampling be redesigned and implemented such that they foster the topic specific self-efficacy beliefs and, thereby, promote higher academic achievement of undergraduate students?

4. How should the overall research design be revised for use in future experiments of the same type implemented in similar contexts?

In order to answer these questions, the present study was comprised of both quantitative and qualitative research components. To complete the quantitative data analysis, I used SPSS version 19 to compute basic descriptive statistics and to determine the coefficients of reliability, G*Power version 3.1.3 to conduct a post hoc power analysis, and Mplus
version 6.1 to analyze the path model. I tested the full path model, which included all hypothesized relations noted in the proposed model as well as two additional interaction terms. I also tested the hypothesis that student self-efficacy beliefs mediated the relationship between the collaborative learning treatment and the academic achievement of the students learning about statistical sampling. Using qualitative research methods, I was able to summarize findings that help explain and add depth to the quantitative data results.

**Descriptive Statistics**

**Survey and Assessment Data**

Table 4 displays basic data describing participants’ survey and assessment results both pre and post instruction. While both the treatment and control groups experienced an increase in self-efficacy beliefs from pretest to posttest, a dependent samples t-test revealed that neither group scored statistically significantly higher at posttest, $t_{control}$ (13) = 1.821, $p > .05$, $t_{treatment}$ (14) = 0.817, $p > .05$. It is important to note, however, that both the treatment and the control groups scored statistically significantly higher on the knowledge measure at posttest (control, $p<0.01$; treatment, $p<0.001$). Further, the treatment group scored higher than the control group at posttest. Differences in pretest knowledge between groups were evident as well, with implications for understanding the effects of the treatment, discussed later.
Table 4

Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th></th>
<th>Treatment</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey or Assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre Unit Self-efficacy</td>
<td>5.470</td>
<td>0.811</td>
<td>5.870</td>
<td>0.536</td>
</tr>
<tr>
<td>Post Unit Self-efficacy</td>
<td>5.770</td>
<td>0.727</td>
<td>5.980</td>
<td>0.223</td>
</tr>
<tr>
<td>Pre Unit Knowledge</td>
<td>50.960</td>
<td>19.400</td>
<td>71.420</td>
<td>13.360</td>
</tr>
<tr>
<td>Post Unit Knowledge</td>
<td>68.270</td>
<td>22.600</td>
<td>83.920</td>
<td>15.050</td>
</tr>
</tbody>
</table>

*Note. Self-efficacy scores could range from 0 to 7. Knowledge scores could range from 0 to 100.*

**Reliability Analysis**

The data from the present study revealed that the coefficient of reliability, $\alpha$, for the self-efficacy survey was 0.919 for the pre survey and 0.852 for the post survey, both above the accepted minimum of 0.800 (Nunnally & Bernstein, 1994). As noted previously, I developed the pre and post knowledge assessment (see Appendix D) based upon the required text for the Basic Statistics course, ensuring that I used the same language as the textbook author. Because these assessments were newly created, no prior reliability data was available for these instruments. The data from the present study revealed that the coefficient of reliability, $\alpha$, for the knowledge assessments were 0.474 for the pre assessment and 0.577 for the post assessment, both well below the accepted minimum of 0.800 (Nunnally & Bernstein, 1994). Prior to the present study, I had never used these pre and post assessments with any group of students. I believe that the fact that these assessments had not been used before, and likely require revision, contributed to the low coefficients of reliability. These low reliabilities had implications for the path analysis, discussed later.
Path Analysis

In order to test the hypothesis that student self-efficacy beliefs mediated the relationship between collaborative learning and academic achievement for these college students studying statistical sampling, as well as to establish the magnitude and strength of the relationships between other variables of interest, I proposed a path model that included pre and post unit student self-efficacy beliefs, pre and post unit student achievement, and instruction type (i.e., direct instruction versus collaborative learning). Using Mplus version 6.1, I was able to determine the relationships between all variables in the hypothesized model and test the hypothesis of mediation. The model estimation terminated normally and, as expected with a just identified model, the model had perfect fit (RMSEA=0.000, CFI/TLI=1.000, SRMR=0.000). Figure 2 shows the original path model, including the strength and statistical significance of each relationship:

Figure 2. Original path model with estimates.
*p < .05, **p < .01, ***p < .001.
Table 5 includes the direct estimates as shown in the model above, as well as the results of a post hoc power analysis. The post hoc power analysis was performed using G*Power version 3.1.3. Also shown is the indirect effect estimate that serves as a measure of mediation of self-efficacy beliefs between collaborative learning and academic achievement in statistical sampling:

Table 5

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Standardized Estimate</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post Unit Knowledge and Pre Unit Knowledge</td>
<td>0.665***</td>
<td>0.993</td>
</tr>
<tr>
<td>Post Unit Knowledge and Pre Self-efficacy</td>
<td>-0.251</td>
<td>0.253</td>
</tr>
<tr>
<td>Post Unit Knowledge and Collaborative Learning</td>
<td>0.061</td>
<td>0.060</td>
</tr>
<tr>
<td>Post Unit Knowledge and Post Unit Self-efficacy</td>
<td>0.220</td>
<td>0.203</td>
</tr>
<tr>
<td>Post Unit Self-efficacy and Pre Unit Knowledge</td>
<td>0.519***</td>
<td>0.858</td>
</tr>
<tr>
<td>Post Unit Self-efficacy and Pre Unit Self-efficacy</td>
<td>0.664***</td>
<td>0.993</td>
</tr>
<tr>
<td>Post Unit Self-efficacy and Collaborative Learning</td>
<td>-0.269*</td>
<td>0.286</td>
</tr>
<tr>
<td>Pre Unit Knowledge and Pre Unit Self-efficacy</td>
<td>0.089</td>
<td>0.073</td>
</tr>
<tr>
<td>Pre Unit Knowledge and Collaborative Learning</td>
<td>0.541***</td>
<td>0.895</td>
</tr>
<tr>
<td>Pre Unit Self-efficacy and Collaborative Learning</td>
<td>0.288</td>
<td>0.324</td>
</tr>
<tr>
<td>Coll. Learning to Post Unit Know. via Post Self-efficacy</td>
<td>-0.059</td>
<td>0.060</td>
</tr>
</tbody>
</table>

* p < .05,  ** p < .01,  *** p < .001

As shown in both the model and table above, the key relationships among collaborative learning, students’ post self-efficacy beliefs, and post unit knowledge did not support my hypotheses. The lack of support for the hypothesized relations led to me to examine whether pretest knowledge or pretest self-efficacy moderated the relationship
between collaborative learning and posttest self-efficacy, and collaborative learning and posttest knowledge. Therefore, I explored other viable path models to determine the best fitting, most parsimonious model for these phenomena.

When one cannot establish statistical significance for a hypothesized mediation path, a common next step in path analysis is to test for interaction, or moderation, effects (Muthen, 2011). Following the methods outlined by Aiken and West (1991), a well-known, straightforward approach to testing interactions (Wen, Marsh, & Hau, 2010), I created two exogenous interaction terms, collaborative learning with pre unit self-efficacy and collaborative learning with pre unit knowledge. Both interaction terms included the categorical variable collaborative learning. Because the categorical variable was binary, no dummy variable coding or unweighted effects coding was necessary (Aiken & West, 1991). Including these interactions required, however, that I mean-center the main effect variables collaborative learning, pre unit self-efficacy, and pre unit knowledge prior to computing the interaction terms. By mean-centering, I was able to reduce collinearity between the interaction and main effect variables. Using Mplus, I then multiplied the mean-centered collaborative learning term with each of the other mean-centered main effects, pre unit self-efficacy and pre unit knowledge, to create the interaction terms.

I tested the new model by regressing both post unit self-efficacy and post unit knowledge on these two interaction terms, as well as on the main effects. Figure 3 shows the results of this analysis:
The relationship between post unit self-efficacy beliefs and the interaction of collaborative learning with pre unit self-efficacy beliefs proved to be the only moderated relationship that was statistically significant. Therefore, I reran the original path model including only this interaction term and the main effects. The chi-square test of model fit was statistically non-significant \( \chi^2(1, N = 27) = 0.228, p = \text{n.s.} \) and the model estimation terminated normally, indicating that the over-identified model had very good fit (RMSEA=0.000, CFI/TLI=1.000, SRMR=0.008). The model displayed in Figure 4 served as the final path model for the present study:

*Figure 3. Interaction path model with estimates.*

\* \( p < .05 \), \** \( p < .01 \), \*** \( p < .001 \).
Table 6 includes the direct estimates as shown in the model above, as well as the results of a post hoc power analysis. Also shown is the indirect effect estimate that served as a measure of mediation of self-efficacy beliefs between collaborative learning and academic achievement in statistical sampling:
### Table 6
Direct and Indirect Effects – Final Path Model

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Standardized Estimate</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post Unit Knowledge and Pre Unit Knowledge</td>
<td>0.665***</td>
<td>0.993</td>
</tr>
<tr>
<td>Post Unit Knowledge and Pre Self-efficacy</td>
<td>-0.251</td>
<td>0.253</td>
</tr>
<tr>
<td>Post Unit Knowledge and Collaborative Learning</td>
<td>0.061</td>
<td>0.060</td>
</tr>
<tr>
<td>Post Unit Knowledge and Post Unit Self-efficacy</td>
<td>0.220</td>
<td>0.203</td>
</tr>
<tr>
<td>Post Unit Self-efficacy and Pre Unit Knowledge</td>
<td>0.457**</td>
<td>0.728</td>
</tr>
<tr>
<td>Post Unit Self-efficacy and Pre Unit Self-efficacy</td>
<td>0.568***</td>
<td>0.931</td>
</tr>
<tr>
<td>Post Unit Self-efficacy and Collaborative Learning</td>
<td>-0.213*</td>
<td>0.193</td>
</tr>
<tr>
<td>Pre Unit Knowledge and Pre Unit Self-efficacy</td>
<td>0.089</td>
<td>0.073</td>
</tr>
<tr>
<td>Pre Unit Knowledge and Collaborative Learning</td>
<td>0.541***</td>
<td>0.895</td>
</tr>
<tr>
<td>Pre Unit Knowledge and Coll. Learning x Pre SE</td>
<td>-0.277</td>
<td>0.302</td>
</tr>
<tr>
<td>Pre Unit Self-efficacy and Collaborative Learning</td>
<td>0.288</td>
<td>0.324</td>
</tr>
<tr>
<td>Pre Unit Self-efficacy and Coll. Learning x Pre SE</td>
<td>-0.429</td>
<td>0.660</td>
</tr>
<tr>
<td>Collaborative Learning and Coll. Learning x Pre SE</td>
<td>0.250</td>
<td>0.252</td>
</tr>
<tr>
<td>Post Unit Self-efficacy and Coll. Learning x Pre SE</td>
<td>-0.224*</td>
<td>0.210</td>
</tr>
<tr>
<td>Coll. Learning to Post Unit Know. via Post Self-efficacy</td>
<td>-0.047</td>
<td>0.056</td>
</tr>
</tbody>
</table>

*p < .05, ** p < .01, ***p < .001

In the final model, one of the principal relationships, the effect of collaborative learning on post unit self-efficacy beliefs, proved to be negative ($\beta$=-0.213) and statistically significant. It is not recommended to interpret main effects in the presence of a statistically significant interaction (Pedhazur, 1997). Therefore, I will interpret this relationship below. Two additional key relationships were positive, but not statistically significant. These are the relationships between collaborative learning and post unit...
knowledge ($\beta=0.061$) and between post unit self-efficacy and post unit knowledge ($\beta=0.220$). A post-hoc power analysis showed that there was insufficient power to establish the statistical significance of these relationships. Therefore, additional study of the relationships between collaborative learning and academic achievement and between post unit self-efficacy and academic achievement for students studying statistical sampling is necessary to better understand the connection between these two phenomena.

**Controls**

Several of the paths in the shown model in Figure 4 are important, but not all are the focal point of the present study. Some of these direct relationships did, however, prove to be both positive and statistically significant. Predictably, pre unit knowledge had a positive, statistically significant relationship with post unit knowledge ($\beta=0.665$). That is, students with more prior knowledge regarding the topics covered in the instructional unit on statistical sampling also exhibited more post unit knowledge on the subject, on average. Pre unit knowledge also had a positive, statistically significant relationship with post unit self-efficacy beliefs ($\beta=0.457$). Students who began this instructional unit with more knowledge of statistical sampling topics exhibited, on average, higher post unit self-efficacy beliefs than students who began the unit with less prior knowledge, controlling for pre unit self-efficacy. Finally, pre unit self-efficacy and post unit self-efficacy shared a positive, statistically significant relationship ($\beta=0.568$). This result aligns with past research evidence that students with high pre unit self-efficacy are likely to have higher post unit self-efficacy than students with low pre unit self-efficacy (Bandura, 1995, 1997; Pajares, 1996; Pajares & Urdan, 2006).
Neither pre nor post unit self-efficacy had a statistically significant effect on post unit knowledge. The relationship between pre unit self-efficacy and post unit knowledge was negative ($\beta=-0.251$). This means that, on average, students with higher self-efficacy prior to the instructional unit tended to perform lower on the post unit knowledge assessment, a result that was not expected. Post unit self-efficacy and post unit knowledge shared a positive relationship ($\beta=0.220$). Thus, students with higher self-efficacy after the instructional unit was complete exhibited higher scores on the post unit knowledge assessment. This result, while not statistically significant, was expected based upon prior empirical research that has shown a positive effect of increased self-efficacy beliefs on academic achievement. Because neither of these results were statistically significant, it is not possible to argue that these findings are different from zero. Nonetheless, the low power associated with these results warrants an examination of these relationships in future research studies.

**Research Question #1: Collaborative Learning and Self-Efficacy**

The relationship between the interaction of collaborative learning with pre unit self-efficacy beliefs and post unit self-efficacy beliefs was negative ($\beta=-0.224$) and statistically significant. Keep in mind that this value represents the treatment effect above and beyond the pre unit measures. Overall, the treatment group scored higher than the control group on both post unit measures (see Table 4). However, this negative value for the interaction indicates that the treatment group gained less than the control group, overall. Recall that the exogenous variables pre unit self-efficacy beliefs and pre unit knowledge were highly correlated with the exogenous variable collaborative learning.
making it difficult to interpret the true effect of the collaborative learning intervention. I will discuss this potential threat to internal validity later in this chapter.

Interactions are often best interpreted by analyzing a graph of the relationship, as shown in Figure 5:

![Figure 5. Relationship between pre and post intervention self-efficacy, split by treatment and control.](image)

The graph indicates that students in both groups saw an increase in self-efficacy over the period of the instructional unit. However, students in the control class saw a statistically significantly greater increase than students in the treatment class. This unexpected result is discussed further in the Discussion chapter. To fully understand the relationships among collaborative learning and self-efficacy beliefs, further investigation of this
interaction and how it relates to the variables in the model should be included in future research.

**Research Question #2: Self-efficacy as a Mediator**

The test for mediation, that is, the effect of the relationship between collaborative learning and post unit statistics knowledge mediated by post unit self-efficacy beliefs also did not prove to be as expected. This indirect effect was negative ($\beta=-0.047$), but not statistically significant. Because this result was not statistically significant, it is not possible to argue that this result is different from zero. However, the relationship between collaborative learning and academic achievement in statistical sampling, mediated by self-efficacy beliefs, is worth investigating in future research studies. These results will be discussed below.

**Relevant Assumptions of Path Analysis**

One of the assumptions of path analysis is perfect reliability of measures (Kline, 2005). Recall that $\alpha$ for the pre self-efficacy survey was 0.919 and for the post survey was 0.852. While these values do not indicate perfect reliability, they do establish that the self-efficacy survey produced reliable data. However, $\alpha$ for the pre knowledge assessment was 0.474 and 0.577 for the post knowledge assessment, well below the accepted minimum of 0.800 (Nunnally & Bernstein, 1994). If both pre and post knowledge assessments had achieved acceptable reliability, I believe the results of the quantitative portion of the present study would have differed greatly. Low reliability results in scores that are invalid (Creswell, 2008), thus adversely affecting the point estimates and standard errors of path analysis. These invalid scores, in turn, hamper the ability to draw meaningful conclusions about the population. With more reliable data,
perhaps I could have established that self-efficacy beliefs do, in fact, mediate the relationship between collaborative learning and academic achievement for college students studying statistical sampling. Because of the unreliability of the data gathered from the knowledge assessment, further study of the relationship among these phenomena is warranted.

A second assumption of path analysis is normal distribution of exogenous and endogenous variables. I looked at the distribution of the data for each of the main effects variables. While not perfectly normally distributed, the distributions of each of the main effects variables were not extremely non-normal and did not give cause for concern. Because of the presence of slight non-normality, I used the estimator MLR when conducting the path analysis. This estimator provides maximum likelihood parameter estimates with standard errors that are robust to non-normality.

**Threats to Internal Validity**

Because the present study did not include random assignment, it was a quasi-experimental intervention study in which I was able to manipulate the independent variable, collaborative learning, by using a treatment versus control design. The results of quasi-experimental studies can help support claims of causality among various phenomena. However, researchers must eliminate threats to internal validity in order to substantiate these claims. Table 7 lists issues that can threaten internal validity and whether or not they apply to the present study:
Table 7

<table>
<thead>
<tr>
<th>Threat</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>History: Events can occur between pretest and posttest that can affect the results of the study</td>
<td>I am not aware of any events that took place between pretest and posttest that would have affected the results of the present study</td>
</tr>
<tr>
<td>Maturation: Effects associated with the passage of time, such as aging, growing tired, etc.</td>
<td>I do not believe that enough time passed between pretest and posttest such that maturation affected the results of the present study</td>
</tr>
<tr>
<td>Instrumentation: Changes in the calibration of instruments or changes in scorers</td>
<td>Neither changes in the calibration of instruments nor scorers occurred during the present study</td>
</tr>
<tr>
<td>Regression: Participants are chosen to participate in a study based upon their</td>
<td>Students were not chosen for each group</td>
</tr>
<tr>
<td>Selection: Groups vary on some characteristic that make them incomparable</td>
<td>Based upon the statistically significant difference between treatment and control groups on pre unit knowledge, it is possible that selection bias threatened the internal validity of the data gathered in the present study</td>
</tr>
<tr>
<td>Attrition: Participants stop participating in a study prior to completion</td>
<td>Due to the large number of students who began the study but did not complete the post unit measures, it is possible that attrition threatened the internal validity of the data gathered in the present study</td>
</tr>
<tr>
<td>Repeated testing: When tested more than once, participants become familiar with the test or become conditioned to being tested</td>
<td>Because there were pre and post unit measures, it is possible that repeated testing threatened the internal validity of the data gathered in the present study</td>
</tr>
</tbody>
</table>
Pretest-treatment interaction: Giving a pretest can change the way in which participants respond to an experimental treatment.

Because both a pre unit knowledge assessment and a pre unit self-efficacy survey were administered to the students, it is possible that pretest-treatment interaction threatened the internal validity of the data gathered in the present study.

The first possible threat to internal validity is selection bias. It is possible that, prior to the study, the treatment and control groups possessed some group characteristic that deemed them incomparable and influenced the outcome results. Based solely on the demographics of the two groups, one would not think this was the case. The treatment and control groups appeared to be very similar regarding gender, race, class, major, and average age. However, the statistically significant difference between treatment and control groups on pre unit knowledge could be due to selection bias and may have affected the outcome of the present study. This difference suggests that the two groups may have been incomparable prior to the beginning of the present study. Recall that the exogenous variables collaborative learning and pre unit knowledge were highly correlated, and that the relationship between collaborative learning and post unit self-efficacy was negative and statistically significant. Because of the strong, statistically significant relationship between the two exogenous variables, the true effect of the collaborative learning intervention was difficult to ascertain. In essence, the pre unit knowledge variable removed some of the effect of the intervention, potentially leading to a biased estimate of the relationship between the intervention variable and post unit self-efficacy. While I cannot definitively state that selection bias is the reason for the negative
relationship between collaborative learning and post unit self-efficacy beliefs, this could be one possible explanation.

A second factor that could have affected the internal validity of the present study is attrition of participants. When people choose to stop participating in a study, for whatever reason, this can lead to a large imbalance in the number of control and treatment group participants who complete all study requirements, thus skewing outcome results. It is possible that participants who drop out of a research study are in some way characteristically different from those who complete all study requirements. Over the course of the present study, both the treatment and control groups experienced a decrease in number of participants. The control group started with 18 students consenting to participate, but ended with 13 participants completing all surveys and assessments. The treatment group began with 23 students who agreed to participate and 14 of those completed all study requirements. Thus, because the final numbers of participants were nearly equal between the control and treatment groups, an imbalance in the number of students who completed all requirements was not an issue in the present study.

Another problem associated with participant attrition is that outcome differences between control and treatment groups may occur because of fundamental character differences between participants who remained in the study and those who did not. That is, participants who chose to complete the study may possess some characteristic(s) that affected the outcome of the study, thus making the research results attributable to that personal quality and not to the treatment itself. The only demographic upon which the dropout participants varied by group was class. In the control group, all three Seniors who consented to participate completed the full study. However, in the treatment group,
none of the five Seniors who consented to participate completed the post unit survey and
assessment. Because this appears to be the only characteristic upon which the two groups
differed, I have no definitive reason to believe that the students who completed all study
requirements were in some way fundamentally different from those students who did not
complete both surveys and assessments. I believe the primary reason for loss of student
participants was student absences on the day the post unit survey and assessments were
distributed and completed.

A third possible threat to internal validity in the present study is repeated testing. When
participants are tested several times over the course of a study, they may become
familiar with the test and make use of correct answers on future tests. Participants may
also become conditioned to know that they are being tested. Both of these issues related
to repeated testing can affect test scores and result in outcomes that are not necessarily
due to the effect of the treatment. In the present study, participants were only tested
twice, before the instructional unit and three weeks later after completion of the
instructional unit. After completion of the pre unit knowledge assessment, the students
and I did not discuss the answers to any of the items on the test. Therefore, they did not
have access to the correct answers for use on the post unit knowledge assessment.
Because students were not tested frequently throughout the period of this research study,
I do not believe that test conditioning was a valid threat to internal validity. I believe that
the results of the present study were unaffected by repeated testing.

The final potential threat to internal validity is pretest-treatment interaction. Recall that both
groups of participants completed a pre unit knowledge assessment and a
pre unit self-efficacy survey. The act of completing these measures could have affected
the way in which students in the treatment group received the collaborative learning intervention. That is, exposure to the pretest measures could have informed students that I was studying self-efficacy and statistics achievement, which could have altered the way in which students behaved during the intervention as well as their performance on the posttest measures. However, I have no evidence to confirm or disconfirm that exposure to the pretest affected participants behavior in the present study.

While several factors could have adversely affected the results of this study, I believe that the only possible threat to the internal validity is selection bias. I believe the statistically significant difference between treatment and control groups with regard to pre unit knowledge may have been a factor that affected my ability to establish positive, statistically significant relationships among collaborative learning, students’ self-efficacy beliefs, and academic achievement of students studying statistical sampling. Next, I discuss the results gathered using qualitative research methods.

**Qualitative Research**

In order to understand the efficacy of the intervention and to appropriately modify the collaborative tasks, as well as to understand how to best revise the foundational learning theory, I used qualitative research methodology. By intently observing the treatment class and by noting the reactions students had to the collaborative tasks and to this type of learning environment, I was able to capture the effect the daily interventions and the class structure had for this group of students. Although 14 students in the treatment class consented to participate in the present study, I observed and recorded notes for all students present in class each day. Because I did not collect identifying information from the students who chose not to participate in the survey and assessment
portions of the study, I was able to gather this qualitative data from all students and from the class as a whole. By observing all student interactions and behaviors, I was able to capture a variety of responses to the intervention and to the collaborative learning environment itself.

I begin with a section summarizing the general results that were common to all class periods. Next, I present the specific occurrences and observational results unique to each class, as well as the daily thoughts captured during each associated reflection period. Finally, I report several important lessons that resulted from my experiences and the changes I made to the foundational learning theory as a result of my daily findings and reflections.

**General Results**

During the six class periods, I noted several unalterable aspects of the classroom environment that affected the students’ experiences during the collaborative learning intervention. The first of these was the architecture of the classroom. As Cohen (1994b) suggested, based upon her research with middle school students, collaborative group work should take place in a classroom that is suitable for this type of learning environment. That is, students should have ample space to work comfortably in groups of four without distraction or interruption by other student groups. Unfortunately, this was not the case in the present study with college students. The classroom in which the intervention class was taught was designed as a small auditorium. There were three rows of fixed, immovable tables, each placed on a riser such that every row was higher than

---

1 This was approved by The University of North Carolina at Chapel Hill Internal Review Board (IRB) under IRB application number 11-1700.
the row in front of it. The rows were slightly curved in arc-like manner. The students’ seats were bolted to the tables and to the floor, and were therefore also stationary. Because of this rather rigid setup, I believed it would have been difficult for students to interact in groups larger than paired dyads. Per the advice of Cohen (1994b) I had originally planned to place students in groups of four, but the architecture of the room did not allow for that type of arrangement. Therefore, students were paired and asked to sit beside their partner during each class period.

Like Cohen’s (1994b) findings with middle school students, forming dyads as opposed to groups of four students proved to be problematic in the college classroom. Because I was forced to pair students instead of assigning them to groups of four as I had planned, student absenteeism had a significant adverse effect upon the collaborative learning treatment. Each class period several students’ partners were absent from class. Therefore, I had to spend time at the beginning of each class re-pairing students. Once I recognized that this would be an ongoing problem, I developed a loose plan to re-group students as necessary for the coming class periods. I made a list of all students within each academic major. Then, within each major, I grouped students by academic performance level. Thus, I devised large groups of students within each discipline and level from which I could choose if necessary to form new dyads.

The classroom architecture also presented a problem for me, the instructor, during classroom observation. As I attempted to move about the room and listen to students’ conversations, trying to ascertain whether or not they were engaging in collaborative learning and to formatively assess their understanding, I found it difficult to physically move among the rows of stationary tables and chairs. The classroom was not designed to
be an active classroom. At times, I had to stand directly over the students to capture their comments and ideas and to listen for the key aspects of collaborative learning such as effective argumentation (Chinn, 2006), negotiation using appropriate academic language (Dillenbourg, 1999; Grover, 2010), and the presence of positive social interdependence (Johnson & Johnson, 1989). I found that my close proximity to the students greatly changed the dynamic of the students’ behaviors and interactions with one another. When I had to stand fairly close to a pair of students to hear their exchange, they would often begin to talk more quietly or stop interacting altogether. Because positive social interdependence allows peer modeling to occur in collaborative learning groups, and peer modeling is one of the primary means of positively impacting students’ self-efficacy beliefs, I knew I needed to have as little effect upon these peer to peer interactions as possible. Realizing this, I tried to fade into the background as much as possible while still capturing the students’ thoughts, ideas, and interactions.

The final general result that was common to the overall teaching experiment was the issue that these students were not my students. The students enrolled in these two classes had become comfortable with and had established a certain level of familiarity and rapport with the primary instructor throughout the course of the semester. Although I had observed these students for three weeks prior to the beginning of this instructional unit and attended class in an attempt to establish a level of comfort with my presence, it was very evident from the first class period of the instructional unit that the students and I had to begin establishing our own relationships. At first, the students did not seem as comfortable with me as they had with their primary instructor. They were not as willing to interact with me, answer my questions, or offer their ideas as I had seen them do with
the primary instructor in the past. In time, however, the students and I began to develop a sense of comfort and ease with one another. I believe that the students and I would have had a very different experience during the present experiment if I had been the primary instructor for this course and had worked with these students for several months prior to implementing the collaborative learning intervention.

**Classroom Observations and Reflections**

**Day 1: Collecting and Using Sample Data Wisely.** On the first day of the instructional unit, 24 of the 32 students registered in the treatment class were present. Because I had pre-arranged the pairs assuming full class attendance, I had to improvise and re-match various students whose partners were absent. Once everyone had a partner, the students began work on the collaborative tasks for the day.

The first section of the instructional unit on statistical sampling focused on how to collect and effectively use sample data. Because the students and I had just finished our discussion of collaborative learning, I asked the students to begin work on this topic without further direction from me. I provided each pair of students with one worksheet (see Appendix F) that contained general directions and two sets of questions to be answered concerning the topic of data collection and use. I also gave each pair of students one small bag of M&M’s candies that were required to complete the first set of questions. The hands-on M&M task required students to think about the concepts of sample, population, data collection, and sampling bias with regard to their bag of M&M’s. Immediately, the students began discussing the general concepts and, specifically, the first task.
As students worked through the first task using the sample of M&M’s candies, I moved throughout the classroom observing and listening as students discussed the ideas of statistical sample, population, data collection, and sampling bias. I specifically listened for any problems the students were having with the content or the way in which the collaborative learning task was presented, as well as for the key characteristics of collaborative learning. The quick acceptance of the collaborative task, engagement on behalf of most students, and thoughtful discourse between partners led me to believe that the students were comfortable with the content itself, the worksheet, and the manipulatives (i.e., M&M’s candies). I heard students accurately use the terms sample, population, and data in their discussions with their partner. Students also exhibited statistical reasoning skills, discussing these concepts in a statistically literate manner. This led me to believe that students understood the concepts of sample, population, and data collection and were beginning to understand how to infer from a sample to a population. I did not, however, hear students discuss the concept of sampling bias during this small group collaborative learning session.

While some of the students were initially uncomfortable interacting with their peers, exhibited by silence or looking down at their desk during the collaborative learning session, most were highly engaged with one another and interested to use the M&M’s to help them consider these topics and ideas. This method of instruction was not the usual mode of instruction in this class, which may have contributed to some students’ unease. After this initial period of transition, however, it seemed that the majority of students embraced this type of learning. I heard comments like, “This is fun!” “We should have candy in every class,” and “This is better than listening to a lecture.” I believed these
positive reactions were due, in part, to the time the students and I spent talking about collaboration and engaging in a collaborative learning task unrelated to the topic of statistical sampling prior to implementation of the collaborative learning intervention. This recommendation from Cohen (1994b) proved to be useful with college students, just as it was for middle school students.

Based upon historical evidence from collaborative learning scholars (see Cohen, 1994b; Erkens et al., 2006; Grover, 2010; Ploetzner et al., 1999), the active, thoughtful discussion and ease with which most students interacted led me to believe that most dyads experienced collaborative learning during this exercise. I also noted that the students were supportive of one another, offering help to their partner if there was confusion or misunderstanding and encouraging each other as they worked through this task. I heard comments like, “No, I think we need to do this” and “Well, but what if we try this?” This indicated to me that the pairs were experiencing positive social interdependence, an atmosphere that can lead to peer modeling (Johnson & Johnson, 1989) and, thus, an increase in students’ self-efficacy (Bandura, 1986).

Once the students seemed to come to conclusion on the M&M problem, I brought the class back together to discuss their experiences working together collaboratively, as well as to discuss any lingering questions, thoughts, or ideas regarding the topics of sample, population, data collection, and sampling bias. Originally, I had planned to have the students spend most of the class period working through all collaborative learning tasks for the day, with little to no interruption from me, before bringing the class back together for a final whole class discussion. Although Cohen (1994b) recommended bringing middle school students together as a class after each task, I thought college
students would have the skills to move through all of the tasks related to the same concepts before re-grouping for a debriefing session. While observing students working through the first M&M task, however, I realized that this was too much to ask of all students in this class, especially those who seemed disengaged during this first collaborative learning exercise. As I walked about the classroom observing the students working together, some stopped me to ask questions while others exhibited confusion or frustration. Therefore, I followed the recommendation of Cohen (1994b) and chose to bring the students back together for whole class discussion after completion of the first set of problems.

The students seemed to appreciate this re-grouping and discussion time, as it appeared to help them determine whether or not they were on the correct thought path and should continue as they had been, or if they needed to modify some of their ideas and conclusions. Students expressed awareness when they realized their understanding was a bit inaccurate. This experience verified that this recommendation made by Cohen (1994b) is not only applicable in the middle school classroom, but is also appropriate for college students studying statistical sampling.

Because I did not hear students discuss the concept of sampling bias in their small groups or during this whole class discussion, I asked the students a series of questions related to the differences between their samples and populations. This question and answer period evolved into whole class discussion and helped students understand the concept of sampling bias. By the end of this discussion, students were using the term ‘bias’ correctly and were able to apply the concept to this task.
The student’s response to these intermittent debriefing periods aligns with the recommendation made by statistical sampling scholars Hodgson and Burke (2000) to include at least one period of whole class discussion to allow students to share and discuss their final thoughts and questions and to exhibit understanding. This debriefing time also allowed me to formatively assess students’ knowledge and understanding. By hearing students use the terms sample, population, data, descriptive statistics, inferential statistics, and bias appropriately, I realized that the students understood the concepts associated with the day’s major concepts. I made a note that the students and I needed to experience these whole class discussions throughout each class period, not solely at the end of class.

After we completed the whole class discussion, I asked the pairs of students to move onto the second task regarding students’ attitudes toward the legalization of marijuana (see Appendix F). I told the students to discuss this scenario using the terminology they developed and used while working on the M&M task, as well as the language we used during the whole class discussion. Again, I circulated throughout the classroom observing, listening, and taking notes.

Besides trying to gain an understanding of how the students received the collaborative learning task, I observed the students to ascertain how the instructional task itself helped or hindered the students’ quest to understand the content and engagement in collaborative learning. I was encouraged to hear students accurately use the terms we had discussed during and after the M&Ms task. Students spoke in a statistically literate manner when discussing the concepts of sample, population, data, descriptive and inferential statistics, making inferences from a sample to a population, and sampling bias.
One issue I noted, however, was that some students were slightly confused by the way I had worded some of the task problems. Several pairs of students stopped me to ask for clarification. I questioned whether or not the students were truly confused or if, because this mode of instruction was foreign to them, they did not trust themselves to work independently of the instructor. As I tried to answer their questions with further questions, I came to believe that their confusion was due more to the latter. When pressed to think about and answer their own questions, they were able to do so. Thus, I made a note and continually encouraged students to work with their partner, struggle with the material, consult other students, and think on their own before they asked questions of me. I believed that encouraging students to challenge themselves to work through the material with their partner, without intervention from me, would push them to productively argue and negotiate, thus leading to collaborative learning (Chinn, 2006; Dillenbourg, 1999; Grover, 2010). I recognized, however, that some of the student’s confusion could be a product of poor wording on my part and therefore made a note to review all task questions and have someone not familiar with this material review the questions for clarity.

By taking their time and working together collaboratively, the students seemed to gain a sound understanding of the topics of data collection and use. My conclusion was supported by the following student exchange:

Student 1: Are we doing this right? Do you think this is right?
Student 2: I think so. Let me see what you wrote.
[Student 2 then reviews what her partner wrote on the worksheet.]
Student 2: Yeah. Remember when we talked about samples in here with [our primary instructor]? That’s the same thing. It’s just part of the whole group.
This dialogue exemplifies one of the key characteristics of collaborative learning, the presence of positive social interdependence (Johnson & Johnson, 1989). In this exchange, Student 2 is helping his or her partner to ensure that Student 1 has a full understanding of what a sample is. Recall that the presence of positive social interdependence is conducive for peer modeling to occur (Johnson & Johnson, 1989) and that peer modeling is one of the primary ways of positively impacting students’ self-efficacy beliefs (Bandura, 1986).

By the end of this task, I heard several students verbally express accurate understanding of these topics. It seemed that being given the time to interact with one another and discuss ideas with their peers helped the students stay on task and maintain an interest in the topics at hand. As student pairs seemed to come to a close on this problem, I brought the class together again to discuss any final questions, ideas, problems, etc. and to formatively assess the students’ understanding of today’s topics. During this whole class discussion, students again exhibited understanding of the topics we had covered on the first day of the instructional unit by accurately using the academic language associated with the concepts of sample versus population, descriptive versus inferential statistics, data collection, and sampling bias. Therefore, I ascertained that both the collaborative learning environment and the task itself helped students understand the material.

The typical mode of instruction in this class, prior to this intervention, was direct, lecture-style instruction. During the three weeks prior to beginning this instructional unit with the students, when I observed the students being taught by the primary instructor, I noted that the majority of students took notes during class. However, when the students
transitioned into collaborative learning mode, I noticed that very few students took notes. Therefore, I decided that at the beginning of each class, I would list on the whiteboard the key topics and terms discussed during the previous class. I had not initially planned to do this, but I believed it would serve as a good reminder of what we covered in the previous class, as well as make clear to the students the key ideas with which they should be familiar. This idea is similar to the debriefing period suggested by Hodgson and Burke (2000). However, I would consider this a ‘rebriefing’ period during which students could individually review or simply note topics from the previous class period.

When I brought the students back together for full class discussions in between tasks and at the end of class, I noticed that it was the same pairs of students who would share their ideas, experiences, thoughts, and questions. Thus, I began to encourage other groups to engage in full class discussion. I was not comfortable calling on specific pairs of students, so I chose to ask for input from “a pair of students who has not yet shared something.” At first, the quieter students were uncomfortable with this. But once they realized that I expected all students to engage and share their ideas, they were more willing to offer input without prompting. Overall, the tasks for the first class seemed to be well received and appeared to help students comprehend the material, while giving them an opportunity to create their own understanding with their peers.

Immediately following the first class period, I participated in the daily period of reflection. During this period, I wrote my thoughts regarding the observations and experiences from the first class. Heeding the advice of several statistical sampling instructors and scholars (see Chang et al., 1992; Fecso et al., 1996; Gelman & Nolan, 2002; Hodgson & Burke, 2000; Kelly, 2010; Mills, 2002; Richardson, 2003; Warton,
2007; Yilmaz, 1996), I tried to develop collaborative tasks that were hands-on, real world activities in order to promote student satisfaction, self-efficacy, and achievement in the area of statistical sampling. During classroom observations, I noted that the students were excited to work with one another and were engaged in the collaborative learning experience. The excitement and high level of engagement on the part of most students helped me realize that the tasks I had developed for the first day were interesting to the students and appeared to be effective in promoting student understanding and collaboration.

In addition, I wrote that I needed to make some changes for the next class period, especially focusing on task wording, the frequency of whole class discussions, and encouraging student interactions. Once I had written my reflection, I transcribed and reviewed the notes I had taken during class observations. Based upon these observations, as well as my reflections, I made the following changes to the tasks and class setup for day two:

1. I re-worded some of the task directions and information to try to make the wording more clear and concise. I had someone unrelated to the present study who is unfamiliar with college level statistics courses review future worksheets for clarity.

2. Following the recommendation from Cohen (1994b), and expanding upon the recommendation from Hodgson and Burke (2000), I made a note to bring the class together for whole class discussion after each individual task (i.e., each set of questions related to one problem), as opposed to waiting until students had completed all tasks on the worksheet.
3. I made a note to encourage the students to struggle with the information and really put thought into the tasks before asking me for help. I also noted that I would suggest that, if they were struggling, to consult with other students in the class before coming to me for help. I hoped this would incite thoughtful discussion and allow students opportunities to effectively negotiate and argue, thus fostering an atmosphere of collaborative learning and peer modeling (Chinn, 2006; Cohen, 1994b; Dillenbourg, 1999; Erkens et al., 2006; Grover, 2010; Johnson & Johnson, 1989; Ploetzner et al., 1999).

4. I developed a list of key terms from the first class that I would list on the whiteboard at the beginning of the second class. I made a note that I would compile this list after each class period and would write the list on the whiteboard at the beginning of the following class.

5. I made a note to begin class on day two with a refresher discussion regarding collaboration. Because some of the students seemed uneasy with this mode of instruction and others were unwilling to share their ideas with the whole class, I believed it was worth ten minutes at the beginning of class to review these expectations and help ease students’ discomfort. Per Cohen’s (1994b) recommendation, I wanted to make sure that students fully understood what collaboration was and developed the skills necessary to effectively collaborate with their partner.

This completed the first of six classes. Once all above mentioned revisions were made to the collaborative tasks and to the class structure, I was ready to move ahead with
the second class period. Below is a summary of the second class period and subsequent period of reflection.

**Day 2: Margin of Error, Confidence Intervals, and Sample Size.** I began class on day two with a brief refresher discussion of collaboration. Per Cohen’s (1994b) recommendation, it is crucial that students understand what collaborative learning is, what is expected of them, and how to develop the skills to collaborate effectively. The students reacted positively to this discussion by adding their own ideas about how to increase collaboration within the student dyads. Student comments like, “My partner and I just jumped right in – it wasn’t so bad” and “Maybe if [students] would start with a conversation about themselves” led me to believe that many students wanted to help their peers become comfortable working within the collaborative learning environment. Most of the students who made suggestions for improvement were those students who seemed comfortable on day one interacting with their peers and sharing their ideas with the class. However, some of the students who were a bit uncomfortable with the collaborative process during the previous class period showed interest in the suggestions offered by their peers. Their body language was positive and receptive, exhibited by making eye contact with the student who was speaking as opposed to looking down at their desks. During this discussion, one of the quiet students expressed his discomfort interacting with his peers in the classroom, but agreed to try to become more engaged with his partner. This experience helped me to understand how important it is to ensure that students at all levels, even college students, are comfortable working in a collaborative learning setting prior to implementing a collaborative learning intervention (Cohen, 1994b).
In the midst of the discussion, I realized that this exchange of ideas also served to provide information about collaborative learning to students who were absent on the first day of class. The students who were not present during our initial discussion were interested to hear what their peers had to say regarding collaborative learning, listening intently and asking questions. I concluded this discussion by encouraging students to struggle with the material and tasks before asking me for help. I also encouraged them to look to other classroom peers for guidance before coming to me for assistance. I hoped that this would help the students engage in deep discussion and encourage negotiation between the students, thus creating and fostering an atmosphere of peer modeling and collaboration (Chinn, 2006; Cohen, 1994b; Dillenbourg, 1999; Erkens et al., 2006; Grover, 2010; Johnson & Johnson, 1989; Ploetzner et al., 1999).

On the second day of the instructional unit, we focused on the concepts of margin of error, confidence intervals, and sample size. Before handing out a worksheet, I listed on the whiteboard the statistical concepts we had discussed and learned during the previous class period. Then, I asked the students to discuss with their partner some of the concepts we covered during the previous class, focusing on data collection and use. I asked them to think about the sample and population of M&Ms from the previous activity and to discuss why the descriptive statistics of their sample differed from those of the full population. The intent of this ‘rebriefing’ period (Hodgson & Burke, 2000) was three-fold. First, I wanted the students to review the information they had learned during the previous class period. Second, I wanted to test Cohen’s (1994b) recommendation and see how the students interacted with one another if they did not have a worksheet to use as a guide. And third, I wanted the students to begin thinking about the idea of margin of
error and an expected difference between sample and population statistics. I meandered throughout the classroom observing, listening, and taking notes regarding the students’ discussions and interactions.

During this time of discussion, most students briefly discussed their thoughts and ideas of the content we had covered during the previous class period and began to talk about the idea of a difference between sample and population values. I heard students engage in short discussions using the terms sample, population, and data. However, without either a worksheet in hand or directions written on the whiteboard to guide them, the students seemed to struggle to engage with one another and to complete the task. Many of the students who participated easily during the first day of class were quieter this class, offering fewer insights and points of discussion. Several pairs of students concluded their discussion much more quickly than I had anticipated. In order for collaborative learning to occur, students need to engage in deep, thoughtful discussion using correct academic language (Chinn, 2006; Cohen, 1994b; Dillenbourg, 1999; Erkens et al., 2006; Grover, 2010; Johnson & Johnson, 1989; Ploetzner et al., 1999). The brief and shallow nature of the student discussions led me to believe that little if any collaborative learning was taking place during this initial exercise. The brief nature of these discussions in which students used few statistical terms from the previous class period also led me to believe that little content was being discussed. Once the students had completed their brief discussions of the material from the previous class period, I brought the students together to share what they had discussed with their partner, a method recommended by Hodgson and Burke (2000).
During this discussion, I asked students how they felt working without a handout. Many expressed unease, stating that they felt, “lost” and “didn’t know where to start”. This way that students reacted to working without some type of written guide supported Cohen’s (1994b) recommendation to provide students with written instructions to guide students through collaborative tasks. Although Cohen’s recommendations were based upon her work with middle school students, I found that college students also struggled without some tangible instructions to use as a guide while working through collaborative learning tasks. This experience also reminded me of an important point that both Gelman and Nolan (2002) and Kelly (2010) noted in their empirical studies regarding statistical sampling. These scholars used handouts to guide students during the discussion of new topics. Because of the unease students exhibited, which I believed was due to the lack of a tangible guide, I made a note to always provide students with some form of explicit, written handout to reference as they worked through the collaborative tasks. It seemed that a written list of questions or instructions helped students more easily navigate the process of completing the tasks with their partner.

Next, I presented a mini-lecture regarding how and why to compute a conservative margin of error. I then distributed one worksheet (see Appendix G) containing three sets of questions to each pair of students. I asked the students to work together to answer only the first set of questions. As the students engaged in collaborative discussion, I moved throughout the classroom observing, listening, and recording notes regarding the students’ ideas, thoughts, questions, collaborations, and discussions. Because this task was rather straightforward, I listened to gain an understanding of the process students used to compute the margin of error for this
problem. The majority of students seemed to have no problem computing the correct value, but the student discussions did not provide good indication of students’ understanding of the concept of margin of error. While observing, I also listened for negotiation among students and the use of appropriate academic language (Dillenbourg, 1999; Grover, 2010), times of productive argumentation (Chinn, 2006), and the presence of positive social interdependence, a component necessary for peer modeling to occur (Johnson & Johnson, 1989).

During the first collaborative learning session of this class period, I noticed that, even after the refresher discussion, a few students remained uneasy collaborating with their peers. As before, these students remained quiet and would not look at their partner. However, I noticed fewer students exhibiting this level of discomfort than during the first class, evidence that some students simply needed extra time to become comfortable with this mode of instruction, a phenomenon that aligns with the findings from research about teaching statistical sampling in a collaborative learning environment (Kelly, 2010). For those students who did engage in collaborative learning, I found that they received the tasks well and that the tasks seemed to help the students learn the concepts. I heard comments like, “Wow – I wouldn’t have thought that” and “That’s interesting.” I also observed students asking other dyads questions when they began to struggle with the tasks or concepts. I overheard one student ask a peer who was not his partner, “Do you think this is what she’s asking?” I even heard one student explain the concept of margin of error to her partner in a different, yet accurate, manner. These observations and comments were encouraging, as I saw them as indications that the students were
beginning to create an atmosphere of positive social interdependence, a key determinant of peer modeling and collaborative learning (Johnson & Johnson, 1989).

Once the pairs of students had completed the first problem, we came back together as a whole class for a period of debriefing (Hodgson & Burke, 2000) to discuss what the students learned from this activity, their ideas, and any lingering questions they had. During this debriefing period, I was able to formatively assess students’ understanding and knowledge by listening for students to exhibit statistical literacy and use statistical reasoning when discussing the concept of margin of error. I heard students accurately explain, in their own words, what the concept of margin of error is and how it relates to the concepts of sample and population. Students noted that the margin of error exists because the sample is, “part of the population, not the whole thing”.

Next, I briefly lectured about how to compute a 95% confidence interval. After this short lecture, I asked the students to work with their partner on the questions in problem two regarding confidence intervals. Again, I walked about the classroom observing, listening, and taking notes regarding the students’ interactions, negotiations with one another, and discussions using appropriate academic language (Dillenbourg, 1999; Grover, 2010). A few students were slightly confused regarding the concept of confidence intervals, stopping me to ask what to do with the margin of error value. I encouraged these students to ask other peers for help and allowed them to struggle with the concept until they seemed to come to understanding. When it seemed that most pairs of students had completed this activity, I brought the students back together to debrief (Hodgson & Burke, 2000) as a whole class to discuss what they had learned, to allow the students to voice any lingering thoughts or questions, and to formatively assess their
understanding of the present topics. I was able to ascertain that the majority of students seemed to understand the concept of confidence intervals based upon their appropriate explanation of the concepts of margin of error, sample, and population, and through their discussions of the relationship between a sample and population. One student even offered an example of a confidence interval regarding public university faculty salaries.

Finally, I briefly lectured on how to calculate desired sample size. Once the students seemed to have an idea of how to compute such values and why it would be of interest to do so, I asked the pairs of students to work through the final set of questions listed as the third problem on the worksheet. Once again I circulated throughout the classroom observing, listening, and recording observations of students’ interactions. Most student dyads exhibited an understanding of how to calculate this value, but none seemed to discuss why this computation was necessary. When the students and I came back together as a whole class for one final discussion to ensure that they had an opportunity to exhibit their understanding, express any last thoughts, and have their remaining questions answered, I asked the students why they should be concerned about determining the appropriate sample size for a research study. They expressed that money, time, and various other resources are typically limited and, therefore, researchers should choose a sample that will provide the necessary data but will not exceed the use of the resources available. I was pleased to discover that, although students had not discussed these ideas during small group collaboration, they understood the reasons behind determining an appropriate sample size.

During the brief lecture periods that occurred between collaborative learning task sessions, I noticed that the students became rather disengaged and distracted. Many
students looked away or were busy with the cell phones, while others just sat and listened, offering no input. Because of the stark contrast between the students’ behavior during brief lecture and their behavior during the collaborative learning sessions, in which most students were highly engaged, I surmised that the students found the collaborative tasks to be more interesting and engaging than lecture. I was not surprised by this result based upon the research literature regarding various methods of teaching statistical sampling to college students (Gelman & Nolan, 2002; Kelly, 2010). Therefore, I made a note to keep the lecture sessions to a minimum.

Similar to the findings of Kelly (2010), the students in the treatment class seemed more comfortable working in a collaborative learning environment on day two than they had during the previous class period. While some students continued to struggle to find their level of comfort with this style of learning, I observed several students making an effort to employ the advice their peers had offered at the beginning of this class period and attempt to engage in the collaborative learning process. I noticed some students who had initially remained quiet and did not want to engage with their partner begin to open up and have discussion with their peers. One initially reserved student even offered feedback to her peers during whole class discussion. After class, I began reflecting upon my observations and experiences of the class period. I noted the effort on the part of most students to try to embrace this type of learning environment, while some students experienced continued unease. I also noted that student interest and engagement appeared drastically different during the collaborative learning sessions and lecture sessions. Finally, I reflected upon the need for some type of written instructions for all collaborative learning tasks.
Considering my reflections along with the notes I took during classroom observations, I made the following changes to the tasks and class format for the subsequent class periods:

1. Following the recommendations from Cohen (1994b), Gelman and Nolan (2002), and Kelly (2010), I made sure to have written directions and instructions, preferably a handout of some kind to distribute to each pair of students, for all collaborative learning tasks.

2. I made a note to continue to encourage students to engage with their peers during the collaborative learning activities.

3. I reviewed the lesson plans for the remainder of the instructional unit and changed areas of lecture to collaborative learning tasks or whole class question and answer sessions where appropriate. Based upon the research literature (see Gelman & Nolan, 2002; Kelly, 2010), I believed this change would keep students more interested in the academic material and help lead to increased confidence and higher academic achievement in statistical sampling (Kelly, 2010).

This completed the second class period of the present study. Once I made the above mentioned changes to the remaining collaborative learning tasks and lesson plans, I was ready to move forward with the next class.

**Day 3: Choosing a Random Sample and Other Sampling Methods (Part One).** On day three of the present study, the treatment class and I covered the concepts associated with developing sampling plans. I began day three by distributing a worksheet (see Appendix H) to each pair of students and asked them to begin work without further direction from me. The students were instructed to work with their partner, to consult
with other pairs of the students in the class if they had a question, and to ask questions of
me only if they came to a true impasse. My intention in asking students to begin work
without input from me was to encourage them to engage in lively, thoughtful discussion
during which time they would have opportunities to negotiate and productively argue
with their partner, thus hopefully fostering peer modeling and collaborative learning
(Chinn, 2006; Dillenbourg, 1999; Grover, 2010). The directions on the worksheet
instructed students to consider developing a research study to better understand the study
habits of students at their university. Each pair of students was asked to develop three
sampling methods to obtain a sample of their university peers from which to gather this
information. They were also asked to explain why they chose each of the specific
sampling methods. Note that the class and I had not yet discussed the various types of
sampling methods prior to this time.

While the students worked to develop their suggested sampling plans, I walked
about the classroom observing, listening, and recording notes about the students’
thoughts, ideas, questions, challenges, negotiations, and productive arguments, as well as
the presence or absence of positive social interdependence; all indications that
collaborative learning is taking place (Chinn, 2006; Cohen, 1994b; Dillenbourg, 1999;
Erkens et al., 2006; Grover, 2010; Johnson & Johnson, 1989; Ploetzner et al., 1999). I
also listened for content understanding and was pleasantly surprised at the level of
creativity exhibited by the students while crafting their sampling plans. Although they
did not know the statistical names of the sampling methods they were discussing (e.g.,
stratified sampling, cluster sampling, etc.), the students were able to devise sampling
plans that included several characteristics of the typical sampling methods discussed in an
introductory statistics class. I heard students talk about choosing, “every 5th student”, discuss how to best use the list of university students (i.e., the sampling frame), and consider using a computer to generate a list of 3000 names. These statistically literate discussions led me to believe that students understood the various single stage sampling plans.

Given the opportunity to create their own ideas with little direction from me, the students quickly became engaged in this activity. They were excited to think about different methods of sampling their peers across campus. I heard statements like, “Oh, but we could do this!” and, “That sounds good, but what if we split the students up by class first?” as well as, “I think that would take too long and it would be too hard to get to everyone.” These comments were encouraging to hear and helped me realize that the students were making a concerted effort to collaborate with their peers to develop what they believed were the best ways to sample the population of university students. I also noted the presence of positive social interdependence, evident from the support and assistance students were offering to one another. When one student was slightly confused or a little off track, I witnessed his partner ask, “What is it that you don’t get? Can I explain this differently?” Observing this led me to believe that peer modeling, one of the phenomena that positively impacts students’ self-efficacy beliefs (Bandura, 1986), was also occurring. I believed that this heightened level of collaboration occurred because the students, over time, were becoming increasingly comfortable with the idea of collaborative learning (Kelly, 2010).

As was the case each day, there were a few students who chose not to participate in the activity with their partner. The students who were paired with a peer who did not
want to participate tried to encourage their partner to engage with them in the task attempting to create an atmosphere of positive social interdependence (Johnson & Johnson, 1989), but typically the non-participating student would sit quietly or offer just a few words during the collaborative learning tasks. The number of students who appeared uncomfortable with this mode of learning or resisted working with their partner was less than during the first two class periods, but this remained a concern for me. I felt as though I needed to try to help all students learn the material as best I could. I began to question my decision from day two to change the majority of class time to collaborative learning or whole class discussion. Thus, I made a note to reflect upon this during after class to determine the best way to proceed.

After all pairs had completed the worksheet, as recommended by Cohen (1994b) and Hodgson and Burke (2000) we came back together as a whole class to debrief and discuss the students’ experiences with this activity. I focused this initial discussion on the students’ experiences with and general feelings about working with a topic about which they had little to no prior knowledge or instruction from me. Overall, most students said they enjoyed the opportunity to create their own ideas and they liked the autonomous nature of the activity, a reaction I anticipated based upon the research of Gelman and Nolan (2002). However, a small minority of students voiced their concerns regarding their ability to engage in and accurately complete a task without prior knowledge or input from the instructor. These included one student who had, up to this point, been fairly disengaged during the collaborative learning tasks, as well as a few students who had previously been active collaborators. Because the third class period
was soon to end, we did not have time to discuss the sampling methods developed by the students and, thus, saved this discussion for the following class.

After the completion of this class period, I reflected upon the events of the day. I reflected heavily upon the issue of a few students’ discomfort with this type of learning environment. It was enlightening to me to realize that not all students wanted to interact with their peers and that some students would rather listen to the instructor lecture. Although Cohen (1994b) warned not to assume that students know how to participate in collaborative working groups, I never considered that students may not want to work collaboratively with their peers. Prior to the present study, based upon my own experience and the research of several college statistics instructors and scholars (Chang et al., 1992; Fecso et al., 1996; Gelman & Nolan, 2002; Hodgson & Burke, 2000; Kelly, 2010; Mills, 2002; Richardson, 2003; Warton, 2007; Yilmaz, 1996), I mistakenly assumed that all college students enjoyed interacting with their peers, be it for social or academic purposes. But I quickly learned that this might not always be true in the classroom. I struggled with this, as I am someone who enjoys collaboration and I am an instructor who prefers to teach in a collaborative learning environment. I understood that not all students learn in the same manner and each student performs differently under different conditions, but it was difficult for me to realize that some students would actually choose not to interact with their peers in a collaborative learning environment. This revelation is something I will remember in my future endeavors as a university professor.

For the present study, however, I had to develop an acceptable compromise for the students and myself such that all students had an equal opportunity to learn the
material, while remaining true to the study of collaborative learning. One of the primary goals of this study was to maximize the efficacy of the collaborative learning treatment, thus providing students the opportunity to learn from their peers as a means of promoting increased self-efficacy beliefs. However, I made what I thought was an acceptable adjustment and modified the previous change I had made on day two regarding the balance of collaborative learning, whole class discussion, and lecture. I re-introduced a few occurrences of whole class discussion. I maintained the change to limit lecture as much as possible, recognizing that this type of instruction is often less stimulating than collaborative learning (Kelly, 2010), but made sure to have an adequate amount of whole class discussion each day in order to provide students with time to debrief about the concepts covered and to ask any lingering questions. Based upon the reflections and observational notes, I made the following changes for subsequent class periods:

1. I reviewed the lesson plans for the remainder of the instructional unit and made sure to have adequate times of whole class discussion to review the topics of the day. This aligns with the recommendation from Hodgson and Burke (2000) to offer students one or more opportunities to debrief during classes that have a large collaborative learning component. This also provided me with an opportunity to formatively assess the students’ understanding and knowledge.

2. I made a note to continue to encourage students to engage with their peers during the collaborative learning activities. This served to offer all students opportunities to use the language of statistical sampling during peer to peer negotiations (Dillenbourg, 1999; Grover, 2010) and productive arguments (Chinn, 2006), and
to create an atmosphere of positive social interdependence (Johnson & Johnson, 1989), all leading to collaborative learning and peer modeling.

This class period and time of reflection completed the third day of the present study. Once I made these changes to the class structure, I was ready to move forward with the next class.

**Day 4: Choosing a Random Sample and Other Sampling Methods (Part Two).** On day four, the students and I continued our whole class discussion regarding the sampling plans they had developed during the previous class period. I asked the students to share their ideas and to discuss the rationale behind the various sampling methods they had designed. This discussion served as a rebriefing period (Hodgson & Burke, 2000) and as the beginning of our work for the day. During this time, I completed a formative assessment, discussed next, of the students understanding of the concepts associated with developing sampling plans. I also took mental notes that I would later add to my written notes.

From the beginning of this discussion, it was evident that the students had discovered nearly all of the basic, well-known statistical sampling methods without prior knowledge of these methods. During our discussion, one pair of students offered the idea of taking the list of names of all currently enrolled students, using a computer to randomly choose 3000 names, and emailing a survey to those 3000 students (i.e., simple random sampling). Another pair of students came up with the idea to separate the student body by class (i.e., Freshmen, Sophomores, Juniors, and Seniors) and choose 750 students from each class to survey (i.e., stratified sampling). A third pair of students discussed standing outside various popular buildings on campus, such as the library, the
main cafeteria, residence halls, etc. and polling people as they exited the buildings (i.e., convenience sampling). Finally, one pair of students developed a method of sampling that involved taking the full list of enrolled students, choosing every fifth student on the list, and emailing a survey to each of the chosen students (i.e., systematic sampling). It was clear from the ideas students shared that they understood the process of most of the well-known sampling methods.

Although none of the students could identify their sampling methods by name, I was pleasantly surprised that they suggested all but one sampling method mentioned in their textbook. I believed that my choice during the previous class period to allow students to begin work on the sampling methods activity without further direction or intervention from me presented students with opportunities to use the language of statistical sampling they had previously learned and to argue, negotiate, and discuss in depth with their peers to determine these sampling methods. As supported by the research literature (see Chinn, 2006; Cohen, 1994b; Dillenbourg, 1999; Erkens et al., 2006; Grover, 2010; Johnson & Johnson, 1989; Ploetzner et al., 1999), this period of struggle while working only with their partner helped the students engage in collaborative learning.

The only method of sampling this group of students did not consider was cluster sampling. Thus, we spent time discussing the process of collecting a cluster sample. Because I observed such disengagement on the part of the students during lecture in the previous class period, and based upon the research of Kelly (2010), who noted a lack of interest on the part of her students when listening to a lecture, I tried to guide the students in whole class discussion to help them comprehend the method of cluster sampling. I
started by asking them to explain the concept of stratified sampling. From there, primarily using questions, I was able to guide the students toward an understanding of cluster sampling in a manner that seemed more engaging to them than lecture.

Once we concluded the discussion of sampling methods, I distributed a new worksheet (see Appendix I) to each pair of students. For this activity, the students were given a project to develop a multistage sampling plan to gather data regarding voter preferences. As students discussed the topic of multistage sampling and developed their plans, I circulated throughout the classroom observing, listening, and taking notes regarding the students’, discussions, levels of negotiation and argumentation, and interactions, especially looking for periods of positive social interdependence between partners.

This task proved to be a bit more difficult than the previous tasks students had completed during this instructional unit. Most students made a concerted effort to begin the activity with their partner by discussing the single stage sampling plans they had previously developed. However, they seemed less confident in their ability to develop a multistage sampling plan than they had while developing several single stage sampling plans during the previous task. As I walked about the classroom and observed the students, some dyads remained rather quiet while others stopped me to ask for further direction to complete the task. I heard several comments and questions like, “I don’t know how to start here” and, “What exactly are you looking for?” as well as, “Is this like the last activity?” These comments led me to believe that this task may have been too complex for collaborative learning to occur without my intervention. Through this observational formative assessment, it quickly became evident that I needed to bring the
class back together to work through this activity together with guidance from me. Thus, I asked the students to stop working on the task with their partner and come back together as a class.

I began the whole class discussion by asking students to share the thoughts and ideas they had created so far, even if they did not believe these ideas and thoughts to be accurate. In my past teaching experiences, I have found this method to be an effective starting point from which to formatively assess students’ understanding and where they are experiencing misconceptions. In this situation, this method of sharing ideas proved to be a useful place to begin. I tried to guide the discussion by asking the students questions rather than directly lecturing to them, as I knew that lecture would prompt them to become quickly disengaged (Kelly, 2010). In essence, this session evolved into whole class (i.e., large group) collaboration.

This experience was very different from the typical whole class question and answer discussions in which we had previously engaged. The students were more willing to offer ideas and collaborate with me and their peers than in the past whole class discussions. One student began by saying, “I don’t know if this is right, but…” and another began with, “I don’t think this is really right….” During this experience, the students recognized that most of their peers were also struggling to complete this task and that they were not alone. The students began to act like a united group, to work together, and offered their feedback, thoughts, and ideas, regardless of correctness or accuracy. Eventually, as a large group, the students began to exhibit understanding of multistage sampling plans. One pair of students offered the plan of first taking a simple random sample of the entire population and then systematically choosing each $k^{th}$ person within
the simple random sample to comprise the final sample. This pair of students believed their plan would create a sample that would not exhaust the resources available to conduct the study. Another dyad suggested that, first, they would stratify the state of North Carolina by region (i.e., mountains, piedmont, and coast). Then, they would cluster each strata by county and choose a simple random sample of counties within each region. These students believed this would provide a good representation of the full population of North Carolina registered voters. During this large group collaborative discussion, I was able to ascertain that most students understood the concepts related to developing a multistage sampling plan and was pleased to hear students use terms like simple random sample, stratified sampling, and cluster sampling in statistically literate ways.

Hearing the students offer their sometimes questionable thoughts and ideas, which eventually transformed into valid methods of developing multistage sampling plans, made me realize that this class was beginning to truly understand how to collaborate, albeit among a large group of their peers as opposed to within dyads. This experience was similar to those reported by Gelman and Nolan (2002) and Kelly (2010). At the beginning of the semester in their statistical sampling courses, these instructors assigned the entire class one large project. The students were separated into smaller working groups, each responsible for a particular portion of the project. In order to provide status reports to their peers, each small student group reported their updates during class time. During this time, other students provided feedback and asked questions of the students reporting. For Gelman and Nolan (2002) and Kelly (2010) these reporting sessions were
very student driven and became time for the class as a whole to collaborate on the overall project.

Through this process, I recognized that this method of instruction, in which the whole class acts as one large collaborative group, can be very rewarding for both the students and the instructor. This was an exciting revelation for me. I realized that, instead of offering lecture as an alternative to small group collaborative learning, I could offer this style of large group collaborative learning, which is more engaging for the students (Kelly, 2010). Because the students interacted differently during whole class discussion than during this large group collaboration, I realized that the students were willing to participate in the large group collaboration only after struggling with their partner during the small group collaborative session. Therefore, I made a note to reflect upon the idea of allowing students to struggle a bit more before intervening and bringing the class back together to work on a task as a large group.

After the completion of this class period, I reflected upon the experiences of the day and completed the post-class reflection. During my reflection, I thought and wrote about this idea of small group collaboration versus large group collaboration. I made the assertion that, for less complex tasks, small group collaborative learning can be an appropriate method of learning for most students. But for more complex topics and tasks, large group collaborative learning may be a better method of instruction. However, I noted that it is imperative to allow students to struggle in a small group prior to engaging in large group collaborative learning, as these experiences of challenge and cognitive disequilibrium helped students organize their thoughts and develop ideas they may not have otherwise. This idea of allowing students to struggle in small groups and only then
coming together as a large group for brief instruction is echoed by Edelson (2001) in his research within the field of Guided Discovery Learning. Based upon the reflections and classroom observations, I made the following notes to help guide the remaining two classes:

1. When developing daily collaborative tasks, if the topic or task seems to be complex, ask students to begin working on the task in small collaborative groups. Be prepared, however, to come together to work on the task in a large collaborative group (i.e., whole class collaboration) if the students seem inordinately challenged by the task. Prior to engaging in large group collaboration, allow the students ample opportunity to struggle with the material in their small collaborative groups. Edelson (2001) has found that this period of struggle and challenge during small group sessions creates a more positive, productive period of large group collaboration.

2. During classroom observations, if it seems that the majority of collaborative student groups are greatly challenged by a topic or task, allow the students ample time to work with the problem in their small collaborative groups. However, when necessary, bring the class back together to work on the task in a large collaborative group. Again, this is supported by the work of Edelson (2001).

This class period and subsequent period of reflection completed the fourth day of the present study. Once I made the above mentioned notes and changes for the remaining class periods, I was ready to move forward with the final two classes.

**Days 5 and 6: Difficulties and Disasters in Sampling and How to Ask Survey Questions.** On days five and six in the treatment class, we focused on difficulties that
can occur when choosing a sample and the challenges of composing good survey questions. On day five, I began class by distributing one worksheet to each pair of students (see Appendix J). The instructions on the worksheet directed the pairs of students to choose a topic they would like to explore and learn more about. Giving students a certain amount of autonomy in choosing their topic can help them develop stronger beliefs in their ability to learn and fully understand the material (Grover, 2010) and gives students in each group a feeling of collective effort, thus inciting interest in the topic (Gelman & Nolan, 2002). The students were instructed to develop one overall question they would like to answer and then write five survey questions to help answer their overall question. The task was designed such that it would incite discussion regarding challenges and mishaps that can occur when sampling a population. Giving students a platform from which to engage in dialogue where they can exhibit their statistical knowledge and literacy can help create an atmosphere of collaborative learning (Cohen, 1994b; Dillenbourg, 1999; Erkens et al., 2006; Grover, 2010; Ploetzner et al., 1999). As the students determined the topic they wanted to explore and developed their five survey questions, I circulated throughout the classroom observing and listening for the accurate use of statistical language, negotiation and productive argumentation among partners, and watching for indications of positive social interdependence.

At first, several pairs of students struggled to determine a topic they wanted to explore and to develop the five survey questions. I observed students sitting more quietly than during the previous class periods and exhibiting less interest in the activity. However, knowing that I had allowed a portion of this class period and the majority of the following class period for large group (i.e., whole class) collaborative learning, and
following the advice of Edelson (2001) regarding Guided Discovery Learning, I chose to allow the students to think about and struggle with the task a bit longer than I might have otherwise. Based on our experiences during the previous class, I believed that allowing the students ample time to determine a topic that interested them would be beneficial to the process of collaborative learning. I recognized that giving the students time to work through their initial hesitations and confusions with their partner would lead to productive discussion, negotiation, argumentation, and positive social interdependence, the key indicators of collaborative learning (Chinn, 2006; Cohen, 1994b; Dillenbourg, 1999; Erkens et al., 2006; Grover, 2010; Johnson & Johnson, 1989; Ploetzner et al., 1999). I also believed that allowing students to struggle with this task could lead to understanding of the concepts associated with developing survey items and the difficulties that can arise when choosing a sample (see Edelson, 2001; Hmelo-Silver, Duncan, & Chinn, 2007). In this scenario, as expected, most students experienced difficulties with engagement and creativity before coming to a place of thought and idea creation. I heard comments like, “I’m not sure where to start” and “I don’t know. What do you want to know more about?” When I allowed the students to spend time thinking about the task, most of the dyads eventually began to engage in discussion, negotiation, and collaboration. I eventually heard students develop topics of interest and use the statistical language they had learned over the course of this instructional unit. I heard students accurately use terms like sample and population, data collection, survey, bias, sampling, and inference. I was pleased to observe students engaged in discussions that were statistically literate and required the use of their statistical reasoning skills. This confirmed for me that, at
times, it is beneficial to allow students to struggle for a bit on their own while working through tasks.

In retrospect, I wondered if I should have allowed more time for students to wrestle with the multistage sampling task from the previous class period. Regardless, I made a note that, for certain students undertaking certain activities, allowing plenty of time for cognitive disequilibrium and challenge can benefit the students in their quest to understand the topic at hand. I also noted that, unfortunately, instructors may not always have enough class time to allow students to progress through this uncomfortable stage of confusion, which often leads to peer modeling and collaborative learning and, thus, increased self-efficacy and understanding. Therefore, as an instructor, it is important to strike an effective balance between offering students the time they need to successfully work through tasks and covering all of the required academic material.

Once it appeared that some of the students had determined a topic they wanted to explore and had developed five survey questions, while other dyads remained challenged by this task, I brought the class back together for a session of large group collaborative learning. Many students were eager to share their ideas and confusions, a discussion that helped to highlight problems that can occur when wording survey questions and sampling from the population. The students who had come to completion on this task offered their ideas to the class while the students who found this task to be a bit challenging listened to their peers and were interested in how they had completed the task. I was encouraged to hear students offer their peers feedback and suggestions for improvement of their survey questions. Even the students who had struggled with this task while working with their partner were engaged in the discussion, both offering ideas and accepting input from their
peers. Similar to the experiences of Gelman and Nolan (2002) and Kelly (2010) in their statistical sampling courses, this class session truly became a collaborative experience, with little intervention on my part. It seemed as if the students were finally comfortable with this style of learning and were beginning to embrace the process. During this large group collaborative session, I was able to formatively assess students’ knowledge and understanding of the concept of multistage sampling. This large group collaborative learning session continued through day six, the final class of the experiment. The student collaborations and interactions during the final class period were similar to those experienced during the large group collaborative learning session the previous class period.

On both days five and six, I engaged in a period of reflection, thinking back upon the experiences of the day. I thought and wrote about the experience of allowing the students to struggle more than normal during the survey question composition task. As an instructor, I noted that this felt somewhat uncomfortable, because I constantly wanted the students to progress and move forward. But I also noted that, for these students, part of progressing was experiencing cognitive dissonance and simply thinking about the task at hand before taking any action. As an instructor, I noted that this realization can be beneficial as I prepare to embark upon my own journey as a faculty member. At times, both students and instructors want to rush through academic material and come to understanding as quickly as possible. However, through this experience, I came to realize that what Edelson (2001) found with his students was true for my students as well. That practicing patience, stepping back and simply observing and listening, and allowing students to experience challenge are often the precursors to deep understanding. Based
upon the reflections and classroom observations, I made recommendations for use during future experiments, as well as on a daily basis in my own classroom:

1. When given the time, allow students to struggle with new material as long as possible and appropriate in order to let them think, process, and experience cognitive disequilibrium. As Edelson (2001) noted, this will be beneficial in helping students come to their own understanding.

2. During large group collaborative learning sessions, as the instructor, intervene as little as possible. If allowed, the students will typically provide very useful ideas and feedback for their peers. This finding aligns with the results reported by Gelman and Nolan (2002) and Kelly (2010).

3. Achieving comfort in a collaborative learning environment takes time, especially if the instructor is not the primary instructor for the course. Three weeks is barely enough time for the instructor to establish a rapport with the students and to help them become comfortable interacting with their peers in this manner. Kelly (2010) noted that, over time, students will become more confident students and will become more comfortable learning in this type of environment and will even take over responsibility for the projects and the class.

Once I had completed the fifth and sixth class periods, the experiment was complete. Two days after the final class, I spent time reflecting on the full teaching experiment and the many lessons I had learned, and revising the foundational learning theory based upon my observations and reflections. Below, I summarize these lessons and modifications.
Lessons and Improvements to Foundational Learning Theory

One of the goals of the present study was to use the data collected during the classroom observations and subsequent reflection periods to make adjustments that help strengthen the foundational learning theory. This occurred during the retrospective analysis phase of the present study. Once the learning theory was modified, it was ready for use in the next iteration of this or other experiments. The present study represents the initial phase of a larger series of interventions.

Recall the following learning theory I proposed prior to implementation of the collaborative learning intervention:

1. With the treatment group of students, I would begin by discussing what the term collaboration meant, how collaborative groups could be effective and productive, and what was expected from each member of a collaborative group. I did not expect that all students would have the skills necessary to engage in collaborative learning and, thus, followed the advice from Cohen (1994b) to discuss these expectations and skills prior to the collaborative learning intervention.

2. Daily, I would use collaborative learning as an instructional method to allow students to work, discuss, and collaborate with their peers. I believed this type of learning environment would give students the freedom to express their thoughts and ideas, feel safe to be wrong, and generate their own comprehension of statistical topics. Allowing for these periods of discussion, negotiation, and argumentation should promote collaborative learning within the student groups (Chinn, 2006; Cohen, 1994b; Dillenbourg, 1999; Erkens et al., 2006; Grover, 2010; Johnson & Johnson, 1989; Ploetzner et al., 1999). I also believed that this
opportunity to collaborate with their peers would help students feel a greater sense of self-efficacy regarding their statistical knowledge and understanding of the concepts associated with statistical sampling, as well as their ability to reason about statistical ideas. This collaborative learning time would serve as the primary means of instruction for the students, with little to no intervention from me, the instructor.

3. I would allow a short period of whole class discussion at the end of each class. The intent of this debriefing period as recommended by Hodgson and Burke (2000) was to ensure that the students understood the material as I expected, to share insights the students had gained while working collaboratively with their peers, and to answer any outstanding questions regarding the statistical topic of the day.

Based upon the lessons I learned through daily observations and reflections during the experiment, I modified the learning theory as follows:

1. Prior to designing a collaborative group intervention, the instructor should consider the architecture of the classroom in which the intervention will take place. The ideal classroom would include movable tables at which four students could sit comfortably and move rather freely. The classroom would also include moveable chairs and ample space for the instructor to move about the classroom and observe students working without impeding student work and interactions.

2. The instructor should not assume that students know how to work together collaboratively, or that all students want to work with their peers in a small group setting. Prior to engaging in tasks related to the academic topic, a whole class
discussion of the meaning of collaboration, how collaborative groups can be effective and productive, and what is expected from each member of a collaborative group should occur. This discussion should include how collaborating with student peers can be beneficial to the learning experience. The instructor should allow students time to express their concerns and apprehensions in regard to working together with their peers, and he or she should try to help alleviate these concerns.

3. On a daily basis, the instructor should use collaborative learning as an instructional method to allow students to work, discuss, learn from, and collaborate with their peers. However, the instructor should be very aware of how the students are accepting this type of learning. Once the instructor has allowed students ample time to struggle with the material and/or this mode of instruction, the instructor may need to intervene with whole class discussion or large group (i.e., whole class) collaboration. During large group collaboration, the instructor should intervene as little as possible. Achieving this balance between allowing students the time to experience sufficient cognitive disequilibrium or discomfort, and helping students attain maximum understanding may vary based upon the context in which the intervention is implemented.

4. During collaborative tasks, it is important that students have some form of written, explicit instructions to follow or questions to answer. Without these written instructions, students may struggle to follow directions and to successfully complete the tasks at hand. Students may also be less inclined to engage with their peers if they do not have a set of written instructions to follow.
5. At the end of each class period, the instructor should bring the class together for a final whole class discussion to ensure that students understand the material as expected, to allow student to share insights they gained while working collaboratively with their peers, and to answer any outstanding questions students may have regarding the topic of the day.

In the Discussion chapter, I will discuss recommendations that led to these modifications.

**Conclusion**

In the present study, I sought to provide an understanding of the relationships among collaborative learning, students’ self-efficacy beliefs, and academic achievement of students studying statistical sampling. Unfortunately, due to various factors presented here and discussed in the next chapter, I was not able to find support for my hypothesis that collaborative, student led, peer groups in the classroom increased the topic specific self-efficacy beliefs of these undergraduate level statistics students. Nor was I able to establish students’ self-efficacy beliefs as the mediating factor between collaborative learning and academic achievement for the participating group of students studying statistical sampling. However, through use of qualitative research, I learned a great deal that has better prepared me for the next phase of this and other experiments.

I learned that the architecture of the classroom plays a very large role in the design and implementation of an intervention experiment. Although Cohen (1994b) mentioned the classroom layout as one of the criteria to consider when developing group work, the results of this study confirmed for me the great influence this factor has on the dynamics of collaborative learning. The classroom size and setup can affect collaborative group size, interactions between the instructor and students, and the
interactions between and among the students themselves, thus fostering or hampering collaborative learning. I also learned that it takes time for students to become comfortable with collaborative learning, especially if this is not their usual mode of instruction (Cohen, 1994b). I had experienced this to a slight degree in the past, but I believe that this experience was different because I had not worked with these students for the portion of the semester leading up to intervention implementation. Similarly, it takes time for the students and the instructor to establish a rapport and a working relationship so that all parties are reasonably comfortable in this type of learning environment. I learned that, when working in small collaborative groups, students should be allowed to appropriately struggle with the tasks and activities independently. However, the instructor must intently observe the students’ interactions and accurately identify the most effective time to intervene. Hmelo-Silver and colleagues (2007) call this ‘just-in-time’ intervention on the part of the instructor. Finally, I learned that, in order to provide the best instruction possible to the greatest number of students, the instructor must achieve the most effective balance of instructional types (i.e., small group collaborative learning, whole class discussion, and large group collaboration) for the students in each class. Because of the overwhelmingly positive student response to large group collaboration, I will discuss recommendations for classroom implementation of this type of Guided Discovery Learning (see Edelson, 2001; Hmelo-Silver et al., 2007; Kuhn, 2007) in the Discussion chapter.

Conducting the present research study using qualitative research methods highlighted various modifications that can help improve future studies of this type. I redesigned the collaborative tasks for this intervention and improved the efficacy of the
tasks for use in this and subsequent experiments. I also gathered data that helped to aptly modify the foundational learning theory such that the new theory can be used as a foundation for future experiments implemented in classroom environments similar in context to that of the present study. I believe that these changes will lead to improved collaborative learning interventions for students who participate in future research studies. These improved interventions are likely to be more efficacious in terms of bolstering peer modeling and its effects upon students’ self-efficacy beliefs. Thus, the information garnered through the use of qualitative research methods will help to improve both the qualitative and quantitative aspects of future collaborative learning intervention research studies. I believe these improvements, along with the lessons I learned and will employ in my future work, can help me and other scholars in the field of education develop more powerful intervention research studies.

In the next chapter, I discuss both the quantitative and qualitative data results of this study and how the qualitative results help to explain and complement the quantitative data results. I also discuss the modifications made to the learning theory and why these changes strengthen the theory for use in the next iteration of this and other intervention experiments. Finally, I offer ideas to conduct an ideal collaborative learning intervention study and make suggestions to improve the next phase of this experiment.
CHAPTER 5
DISCUSSION

It has been well established that increased self-efficacy beliefs, whether generally academic or domain, topic, or task specific, are positively related to academic performance (see Chemers et al., 2001; Finney & Schraw, 2003; Pajares & Miller, 1997). It has also been shown that collaborative learning is a more successful method of promoting academic achievement than individual or competitive learning (Johnson & Johnson, 1989; Johnson et al., 1981). During collaborative learning exercises, students are able to observe their peers working to successfully complete academic tasks. The peer modeling aspect of collaborative learning can, in turn, increase students’ self-efficacy beliefs, leading to greater academic success (Bandura, 1995). Understanding new statistical ideas and topics can be a challenge for undergraduate students (Garfield, 1995). Statistical sampling is an area of statistics that students often find to be overly abstract and complex (Nguyen, 2005; Yilmaz, 1996). For example, students struggle to understand the concepts of sample and population and the relationship between the two (Saldanha & Thompson, 2002; Watson & Moritz, 2000), and often overlook various aspects that are important when developing and administering sample surveys (Schwartz & Goldman, 1996). In an effort to help undergraduate statistics students achieve, instructors need to know more about how collaborative learning in the undergraduate statistics classroom affects the self-efficacy beliefs, and thus academic achievement, of students studying statistical sampling.
In the present study, I attempted to establish students’ topic specific self-efficacy beliefs as the mechanism by which collaborative learning positively affects the academic performance of college students studying statistical sampling. In order to support the claims of causality among these phenomena, I conducted a quasi-experimental intervention study. To ensure that the intervention was administered according to plan, as well as to offer suggestions to researchers and practitioners who choose to conduct future studies of this type, I employed qualitative research methods to study the implementation and efficacy of the collaborative learning intervention. Using these qualitative methods allowed me to determine how the daily intervention tasks should be revised based upon my observations and reflections. The use of this methodology also helped me make appropriate modifications to the foundational learning theory in order to strengthen the theory for use in future research experiments.

In order to understand the relationships among collaborative learning, self-efficacy beliefs, and academic achievement of students studying statistical sampling, as well as to make refinements to the intervention tasks and the foundational learning theory, I sought answers to the following four research questions:

1. Does working in small, collaborative, student led, peer groups in the classroom increase the topic specific self-efficacy beliefs of undergraduate level statistics students?

2. Do topic specific student self-efficacy beliefs mediate the relationship between collaborative learning and the academic achievement of undergraduate level statistics students studying statistical sampling?
3. How should the collaborative tasks intended to help undergraduate students learn topics related to statistical sampling be redesigned and implemented such that they foster the topic specific self-efficacy beliefs and, thereby, promote higher academic achievement of undergraduate students?

4. How should the overall research design be revised for use in future experiments of the same type implemented in similar contexts?

The effect of collaborative learning upon post unit self-efficacy beliefs did not prove to be as I had hypothesized. That is, I was not able to establish that working in small, collaborative, student led, peer groups in the classroom increased the topic specific self-efficacy beliefs of these undergraduate statistics students. I was also unable to empirically support the hypothesis that topic specific student self-efficacy beliefs mediated the relationship between collaborative learning and academic achievement for these students studying statistical sampling. Based upon the observations of and experiences with the students who participated in the intervention, I present several reasons why I believe I was not able to provide evidence to support these two research hypotheses.

**Research Question #1: Collaborative Learning and Self-efficacy Beliefs**

The first research question of the present study focused on the direct relationship between collaborative learning and students’ self-efficacy beliefs. I hypothesized that undergraduate statistics students’ topic specific self-efficacy beliefs would increase as a result of working together with their peers in small, student led, collaborative groups within the classroom. I was unable to provide support for this hypothesis. I believe there are several factors that affected the ability to establish a positive, statistically significant
relationship between collaborative learning and self-efficacy beliefs for the students who participated in this study. These include possible selection bias due to a high correlation among exogenous variables; small sample size; various barriers to students’ experience of peer modeling, including classroom architecture, group size, and student attendance; the novelty of this type of learning environment for most of the students in the treatment group; the short time period of intervention implementation, which contributed to the challenge of establishing a comfortable working relationship among the students and between the students and myself; and a possible upward social comparison effect.

Selection Bias

Selection bias, or outcome differences between the treatment and control groups due to prior differences between groups, may have threatened the internal validity of the present study, thus affecting the results. After data collection, it was established that the treatment and control group participants statistically significantly differed on their levels of pre unit knowledge. This difference could be due to selection bias and may have affected the outcome of the present study. While I cannot definitively say that selection bias is the reason for the negative, statistically significant relationship between collaborative learning and post unit self-efficacy beliefs, a result that was unexpected, this is one possible explanation.

Sample Size

It is well known that data gathered from a large sample provides better estimates of population values than data gathered from a small sample (Connolly & Sluckin, 1953; Cohen, 1988; Cohen & Lea, 2004). The smaller the sample size, the more inaccurate the methods of computing population estimates become (Connolly & Sluckin, 1953).
Sample size and statistical power are similarly related (Cohen, 1988). That is, as the sample size increases, so does, “the probability of detecting the phenomenon under test” (Cohen, 1988, p. 7). I believe that the small number of students who participated in the present study contributed to the inability to demonstrate that collaborative learning in the classroom helps students increase their topic specific self-efficacy beliefs.

Prior to conducting the present study, I ran a power analysis assuming full participation (i.e., 30 students per group). The a priori power analysis revealed that, at any effect size (i.e., small, medium, or large), I would have difficulty establishing the statistical significance of the relationship between collaborative learning and students’ self-efficacy beliefs. Recall that only 13 of 29 students enrolled in the control class and 14 of the 32 students enrolled in the treatment class completed all pre and post surveys and assessments. Thus, full data from only 27 students, even fewer than expected, were available with which to conduct the quantitative analyses. Because I was limited to voluntary participation of students enrolled in two sections of a Basic Statistics course, I was unable to control the number of students who participated in this study. Considering the low number of participants, it is not surprising that I was unable to find statistical significance to support my hypothesis. I believe that access to a larger student population and consent from a greater number of participants would have increased the statistical power of the path estimate between collaborative learning in the classroom and the self-efficacy beliefs of undergraduate students studying statistical sampling.

**Barriers to Peer Modeling**

Bandura (1995) proposed that vicarious experience, or peer modeling, is one of the four primary mechanisms by which students can positively affect others’ self-efficacy
beliefs. The process of observing peer models work diligently to reach high academic achievement has been shown to have a significant positive effect upon the academic achievement of the observing student (Cohen, 1994a; Johnson & Johnson, 1989; Pajares 1996). Schunk and Hanson (1985) also found that observing high achieving peer models had a statistically significant positive effect on the topic specific self-efficacy beliefs of elementary school children. Peer modeling is one of the benefits of employing collaborative learning in the classroom (Johnson & Johnson, 1989). There are several aspects of the present study that may have had an adverse effect on the peer modeling experiences of students in the intervention class, thus affecting the ability to establish a positive, statistically significant relationship between collaborative learning and students’ self-efficacy beliefs for statistical sampling. Many of these disadvantageous aspects were a result of the architectural design of the classroom in which the intervention occurred.

**Classroom architecture.** Instructors who plan to implement small group collaborative learning in their classrooms are encouraged to consider several criteria prior to implementation (Cohen, 1994b). One of these considerations that can affect various aspects of an intervention study is the layout or architectural design of the classroom. Cohen (1994b) recommended that collaborative learning occur in a classroom that allows students to be arranged in groups such that all students in the group can easily interact with one another, preferably seated in a circle. She also suggested that the classroom be large enough such that teachers can position the student groups as far as possible from one another to limit noise disruptions from other groups. She warned that, “lack of work space can result in disengagement and general failure of the projects” (Cohen, 1994b, p. 76).
Due to the layout of the classroom in which the present intervention study occurred, a small auditorium containing immovable tables and chairs, students were not able to sit face to face while working together in collaborative pairs. Instead, students had to sit next to one another and position themselves in sometimes awkward ways in order to discuss and interact. Similar to Cohen’s (1994b) findings with middle school students, this constrained seating arrangement inhibited these college students’ interactions and tainted the collaborative learning experience for the students, affecting their ability to truly engage with one another and to create an atmosphere of peer modeling, thus hampering opportunities to positively affect their statistical sampling self-efficacy beliefs.

The architectural design of the classroom in which the intervention occurred also affected my experience, which, in turn, affected students’ interactions. As the instructor and sole researcher, I found it difficult to freely move about the classroom to observe students’ interactions. Due to a lack of space and the close arrangement of immovable classroom furniture, I often found myself standing so close to student dyads in order to listen and observe that I almost became part of the group. I believe that these physically close observations changed the dynamics of peer interaction within the dyads. That is, I believe students would have interacted more openly with one another had I been able to observe from a greater distance. This would have allowed students to experience collaborative learning in an atmosphere of peer modeling, thus providing opportunities to positively affect self-efficacy beliefs.

**Group Size.** Through her research on small, student led, collaborative group work with middle school students, Cohen (1994a) found that four to five students is typically
The ideal group size. She noted, however, that the tasks themselves may determine how many students should be placed in each group. For example, when the instructor wishes to foster an environment in which student groups engage in lively, academic discussion, Cohen noted that groups larger than five can create a scenario where some students may be left out, while groups smaller than four may not be large enough to provide sufficient divergent viewpoints and ideas to incite productive discourse and provide ample peer models.

The arrangement of students in dyads as opposed to groups of four, a decision dictated by the architectural design of the classroom, affected the ability to establish that collaborative learning in the college classroom and students’ topic specific self-efficacy beliefs share a positive, statistically significant relationship. This experience confirmed that Cohen’s (1994b) findings with middle school students also apply in the college classroom. Group size adversely affected the results of the present study in two ways. First, I believe that arranging students in dyads severely diminished the breadth and depth of the discussions and interactions that occurred during the collaborative learning tasks. It has been shown that productive, small group student discussion can be an effective way for students to develop deep understanding of a topic (Johnson & Johnson, 1989). In a statistics classroom, these student led discussions can act as a platform for students to exhibit statistical reasoning skills, a tool necessary for developing statistical literacy (Ben-Zvi, 2005; Ben-Zvi & Garfield, 2005; delMas et al., 2007; GAISE College Report, 2005; Garfield, 1995; Garfield, 2005). Because, during small group collaborative learning sessions, students were engaged in discussion with only one peer, they did not experience ample discourse from which to develop these statistical reasoning skills.
These sparse discussions likely affected the efficacy of the intervention, thus making it unlikely that I would find a positive, statistically significant relationship between collaborative learning and the self-efficacy beliefs of students studying statistical sampling.

Second, I believe that arranging students in dyads did not expose students to a sufficient number of differing peer perspectives and potential peer models, thus diminishing students’ opportunities to interact with peer models that could positively impact their topic specific self-efficacy beliefs. Recall the suggestion from Wang and Lin (2007) that practitioners place at least one student with high topic specific self-efficacy beliefs within each collaborative learning group. If students had been organized in groups of four, the likelihood that at least one student in each group was academically advanced enough to serve as a model for his or her peers would have been greatly increased. Therefore, arranging students in dyads adversely affected the ability to establish support for the hypothesis that peers learning together in small collaborative groups within the classroom increases students’ self-efficacy for statistical sampling.

**Student attendance.** In order to fully engage in collaborative learning and have opportunities to learn from one another, students were dependent upon their partner being present in class each day. When a student’s original partner was absent from class, they were reassigned to a new partner for the day, an event that occurred regularly. This caused inconsistencies in partnership for several students. Due to student absenteeism, some students had up to four different individual partners throughout the six class periods. Each time a student was assigned a new partner, he or she had to reestablish familiarity and comfort, a process that could take several class periods for certain
students. While it has been established that peer modeling is one method of helping students improve their self-efficacy beliefs (Bandura, 1995), it is unclear how changing partners during an intervention, thus creating inconsistencies in peer modeling, influences students’ self-efficacy beliefs. I believe these inconsistencies in partnership, which, for several students led to the need to continually reestablish a comfortable rapport between partners, was a contributing factor to the inability to establish a positive, statistically significant relationship between collaborative learning and the self-efficacy beliefs of students studying statistical sampling.

Additional Obstacles

Other factors, beyond the problems associated with the architectural design of the classroom, negatively affected the outcomes of the present study.

New learning environment. Johnson and Johnson (1989) posited that collaborative learning promotes greater academic success than individual or competitive learning environments. Cohen (1994b), however, warned that instructors cannot assume that students automatically know how to collaborate with their peers in the classroom. She suggested that instructors must take the time to teach their students the social skills necessary to interact with their peers and to discuss the expectations of collaborative group members. Light, Light, Nesbitt and Harnad (2000) agreed, stating, “interactivity rarely occurs spontaneously, even in a well-supported environment” (p. 200).

Recall that the usual mode of instruction in both the control and treatment classrooms prior to the intervention was primarily lecture, whole class question and answer, and whole class discussion. During the three weeks that I observed both classes prior to teaching the instructional unit on statistical sampling, I did not observe the
students working together in small groups to complete tasks related to the academic material. Thus, the collaborative learning environment that I introduced in the treatment class was new in this course, for this group of students. Because this was the case, I followed Cohen’s (1994b) advice and spent time during the first two class periods teaching the treatment group of students about collaboration, discussing with them their definitions of the term, and setting expectations that I believed would help foster an atmosphere of collaboration.

During classroom observations, I noticed that it took several students a good bit of time, often several class periods, to become comfortable engaging in this type of learning environment. In fact, some students never seemed at ease working with their peers, creating their own understanding, and working independently of the instructor. Based upon both my prior experience with collaborative learning and my observations during the present study, as well as the recommendation from Cohen (1994b) based upon her research with middle school students, I believe that this style of learning requires a great deal of time and practice for college students to achieve maximum benefit. Because the collaborative learning environment was completely new to the students, I believe that three weeks was not sufficient time to fully engage in effective collaboration and to become comfortable in this new environment such that the students could begin to benefit from this style of learning. Just as Light and colleagues (2000) stated, some of the students did not spontaneously interact. I believe that the length of time spent working in a collaborative learning environment hindered students’ opportunities to engage in effective collaboration, thus making it unlikely that students’ specific self-efficacy beliefs for statistical sampling would increase over such a short period. For future interventions,
I plan to spend a full semester working with students to help them become comfortable in this type of learning environment and providing them the opportunity to experience full collaboration with their peers.

**Interim instructor.** Purvis and Garvey (1993) attested that a substitute teacher is someone who provides students with an experience “equal” to that which is typically provided by the primary instructor. Yet Lassmann (2001) questioned whether parents, other teachers, and students consider substitute teachers to be equal to the primary teacher. She suggested that the term “interim” teacher may be a better way to define teachers who assume the daily responsibilities and duties of the regular classroom teacher. Nidds and McGerald (1994) interviewed a group of substitute teachers in an effort to elucidate the primary issues substitute teachers face in the classroom. They revealed that the most difficult problem substitute teachers encounter is a lack of knowledge of the environment in which they are working. That is, substitute teachers often do not know the individual students, the typical routine of the classroom, or the school’s policies and procedures. This lack of knowledge leads to behaviors from students that are typically not exhibited when the primary instructor is present in the classroom. That is, students simply act differently when a substitute teacher is responsible for the daily classroom activities.

Both the primary professor for these two Basic Statistics courses and I made it clear to the students in both classes that I had full responsibility for the classes during the three weeks of the instructional period. However, I believe that the fact that I was not the primary instructor for this course, and thus the students viewed me as a substitute teacher, played a large role in the inability to effectively implement the collaborative learning
treatment. By attending class for three weeks prior to the instructional period, I tried to establish a sense of familiarity and comfort with and for the students. I did not, however, interact with the students during this observational period, thus limiting the students’ familiarity with me.

Just as students require time and consistency to establish a rapport and level of comfort with their peers, it also takes time for an instructor to establish familiarity with and to gain the respect and trust of his or her students. I am unsure if the students in the intervention class trusted that this new mode of instruction, this collaborative learning environment that I thrust upon them, would help them understand the material. While most students seemed to be willing participants, at the end of the three week instructional period there remained a few students who remained uneasy learning in this type of environment. I believe that if the students and I had the opportunity to first establish a working relationship prior to implementation of the intervention, a greater number of students would have been willing to engage in the collaborative learning process. And, those students who did embrace the collaborative learning environment would have been able and willing to more fully engage in collaboration than they did during our three weeks together. Thus, if the students had been comfortable with me as the instructor, I believe they would have felt more at ease working within a collaborative learning environment and, hence, would have had greater opportunity to engage in experiences that could have positively influenced their statistical sampling self-efficacy beliefs.

**Upward social comparison.** In the present study, the self-efficacy beliefs of students in the treatment group did not increase as much as I had anticipated. I was unable to show that undergraduate statistics students’ topic specific self-efficacy beliefs
statistically significantly increased as a result of working together with their peers in small, student led, collaborative groups within the classroom. I believe one aspect of collaborative learning that may have influenced this result is the phenomenon of upward social comparison. An upward social comparison effect occurs when a student compares him or herself to a peer who exhibits greater competence (Carmona, Buunk, Dijkstra, & Peiro, 2008; Miyake & Matsuda, 2002; Vrugt & Koenis, 2002). The effect of this comparison could be an initial decrease in self-efficacy beliefs which, in turn, can decrease academic performance (Carmona, Buunk, Dijkstra, & Peiro, 2008).

Historically, it has been assumed that upward social comparisons, or comparing oneself to someone who is perceived to perform at a higher level, can initially decrease students’ self-efficacy, confidence, self-esteem, and various other feelings about themselves and their performance (Carmona, Buunk, Dijkstra, & Peiro, 2008; The Encyclopedia of Social Psychology, 2007). However there is evidence to the contrary, which states that both upward and downward social comparisons can have either positive or negative effects upon students’ affect (see Buunk, Collins, Taylor, VanYperen, &Dakof, 1990; Johnson & Stapel, 2007; Taylor & Lobel, 1989). Thus, the effect of social comparisons remains a controversial topic. Regardless, it may be that students in the treatment group of the present study initially experienced a decrease in self-efficacy as a result of upward social comparison during the collaborative intervention. That is, the act of comparing themselves to a peer who outperformed them could have served to decrease their confidence in their abilities prior to becoming comfortable in this new learning environment and eventually experiencing an increase in self-efficacy. This may
explain why the treatment group of students showed a smaller pretest to posttest gain in self-efficacy than the control group.

As noted, I believe that the short time period of this collaborative learning intervention impeded my ability to establish self-efficacy as a mediating factor between collaborative learning and academic achievement. If some students in the treatment group experienced an initial decline in self-efficacy before experiencing increased self-efficacy, the relatively short amount of time for this intervention could have been a negative factor. That is, the intervention period may not have been long enough to fully overcome the initial decrease in self-efficacy and to realize the full positive effect of collaborative learning upon students’ self-efficacy beliefs. Future intervention studies of this type should be implemented over the period of at least a full semester in order to allow students to become comfortable with this type of learning environment and to overcome any initial setbacks in self-efficacy.

**Aspects of a Strong Intervention Study**

While no intervention could ever be perfectly developed and implemented, Hsieh and colleagues (2005) proposed what they believed are features of a strong educational intervention research study. These included:

1. Random assignment, or the ability to assign students to groups
2. If unable to randomly assign students to groups, the ability to minimize threats to internal validity
3. Treatment integrity
4. Using multiple outcome measures
Because the students who participated in the present study self-selected into each of the course sections I taught, I was not able to randomly assign students to the control and treatment groups. However, I believe that I was able to successfully minimize the potential threats to the internal validity of this intervention study (see “Threats to Internal Validity” above). Based upon the recommendations from Hseih and colleagues (2005), I believe that these minimizations helped to strengthen this intervention study.

In the present study, I faced several obstacles while trying to implement an effective collaborative learning intervention. These included the architecture of the classroom, the size of student groups, the small sample size, and the length of time of the intervention. Despite these challenges, I did my best to implement the intervention properly and to understand the efficacy of the collaborative learning treatment. Using qualitative research methods, I was able to identify aspects of the present study that should be improved during the intervention, as well as modifications to be made for future studies of this type. I believe these data regarding treatment integrity helped to provide a more effective collaborative learning intervention for the students in the treatment class.

In order to understand the relationships among collaborative learning, students’ self-efficacy beliefs, and academic achievement, I used multiple quantitative and qualitative measures. I administered a quantitative survey to assess students’ pre and post unit self-efficacy beliefs and a quantitative knowledge assessment to assess students’ pre and post unit understanding of the concepts related to statistical sampling. Qualitatively, I observed and took notes while students worked together in collaborative dyads. My observations focused upon students’ interactions, behaviors, verbal thoughts and ideas,
and non-verbal language regarding their reception to the collaborative learning tasks. I also reflected upon my experiences after each class period. I used the notes from my observations and reflections to help improve the efficacy of the collaborative learning treatment. I believe that the use of multiple quantitative and qualitative measures helped to strengthen this collaborative learning intervention.

Hsieh and colleagues (2005) noted several aspects of a strong educational intervention research study. With the exception of random assignment, I believe the present intervention study met these criteria. Although I was unable to establish that undergraduate statistics students working together in a collaborative learning environment experienced a statistically significant increase in self-efficacy beliefs, I believe I implemented the best intervention study possible given the various obstacles I faced.

Conclusion

There are many factors that I believe played a role in the inability to implement an effective treatment, and subsequently establish a positive, statistically significant relationship between collaborative learning and the self-efficacy beliefs of students studying statistical sampling. These factors included possible selection bias; the small sample size; the architectural design of the classroom which affected students’ interactions, assigned group size, the peer modeling experiences of students, and contributed to inconsistencies in partnership between students; the novelty of this type of learning environment; and the short time period of intervention implementation, which contributed to the challenge of establishing a level of comfort and rapport among the students and between the students and myself, and may have interfered with students’
opportunities to experience the full effect of collaborative learning upon self-efficacy beliefs. Next, I discuss plausible reasons why I was not able to establish students’ topic specific self-efficacy beliefs as a factor that mediated the relationship between collaborative learning and academic achievement of students studying statistical sampling in the present study.

**Research Question #2: The Mediation Effect of Students’ Self-efficacy Beliefs**

It has been well established that collaborative learning leads to greater academic achievement than individual or competitive learning (Johnson & Johnson, 1989; Johnson et al., 1981). The second research question of the present study focused on whether or not students’ topic specific self-efficacy beliefs mediated the relationship between collaborative learning and academic achievement for students enrolled in an undergraduate introductory statistics course studying statistical sampling. I tested this hypothesized mediation effect as part of the path model analysis. Due to several factors, I was not able to establish students’ self-efficacy beliefs as a mediating factor between collaborative learning and academic achievement for the students who participated in this study. Using quantitative and qualitative research methods, I was able to determine various issues I believe contributed to the inability to establish this statistically significant mediation relationship. These include possible selection bias due to strong correlation among exogenous variables; small sample size; grouping students in dyads, which contributed to inconsistent partnership between students and less opportunity for peer modeling to occur; the social dynamics of collaborative learning groups; and the short length of time I had to spend with the students.
Selection Bias

As mentioned, I believe the only plausible threat to internal validity of the present study was selection bias. If selection bias was truly a factor, this could have affected my ability to establish self-efficacy beliefs as a mediating factor between collaborative learning and academic achievement. Because the treatment group was statistically significantly different from the control group regarding pre unit knowledge, there is reason to believe that these groups differed in some characteristic manner that deemed them incomparable and negatively affected the results of the present study. While I cannot definitively say that selection bias is the reason that I am not able to establish self-efficacy as a mediating factor between collaborative learning and academic achievement, it is one possible explanation.

Sample Size

As mentioned previously, the sample size for the present study was smaller than anticipated. I believe the small sample size was one of the primary factors that contributed to the inability to determine that students’ topic specific self-efficacy beliefs mediated the relationship between collaborative learning and academic achievement for these students studying statistical sampling. I believe that, given a larger sample of students, the test for mediation would have been more powerful, and I would have been able to provide positive, statistically significant support for the mediation hypothesis.

Student Dyads

Peer modeling, a characteristic of collaborative learning (Johnson & Johnson, 1989), is one mechanism by which students can increase their self-efficacy beliefs (Bandura, 1995). In order to foster effective collaboration and create an atmosphere of
peer modeling, students should be arranged in groups of four to five if the goal is to incite lively, diverse discussion (Cohen, 1994b). Due to the architectural setup of the classroom, I was restricted to arranging students in pairs, as opposed to groups of four as originally planned. Because of this arrangement, students were not exposed to a variety of peers within their group, thus decreasing the likelihood of interacting with at least one student advanced enough to serve as a peer model. I believe the lack of peer exposure adversely affected the ability to establish students’ topic specific self-efficacy beliefs as a mediating factor in the relationship between collaborative learning and academic achievement in statistical sampling.

The dyadic structure of the collaborative groups negatively affected the outcome of the present study in another way. If students in the treatment class were given the opportunity to work with the same group members for the full six class periods, the partnerships would have likely evolved into familiar, comfortable, effective, collaborative relationships. Each time a student was assigned a new partner, I observed both students begin by introducing themselves, trying to establish a relationship and process of working together, and generally trying to find a level of comfort with one another. Because many students experienced this initial meeting process several times, students’ abilities to fully engage in collaborative learning and to create an atmosphere of peer modeling were impeded.

If students had been arranged in groups of four as originally planned, student absences may not have presented the issues produced by being assigned only one partner. In larger groups, where the absence of one group member would not have been devastating to the function of the group, students would not have been reassigned new
partners and would likely have been able to establish a comfortable working relationship that would have grown over the course of the intervention period. I believe this constant change in partnership did not allow students to fully engage in collaborative learning and establish an atmosphere of consistent peer modeling, thus contributing to the inability to establish students’ self-efficacy beliefs as a mediating factor between collaborative learning and academic achievement in statistical sampling for this group of students.

Social Dynamics of Collaborative Learning

It has been established that collaborative learning can increase students’ academic achievement (Cohen, 1994b; Johnson & Johnson, 1989), student course satisfaction, and student confidence (Kelly, 2010). However, there is evidence that not all students experience these benefits as a result of working with their peers in small, collaborative groups in the classroom (Webb & Mastergeorge, 2003). This is especially true for students who struggle with various concepts or academic material. This may be due, in part, to the ways in which students interact with their peers in collaborative learning groups.

Students learn from their peers in a variety of ways, including giving and receiving help, sharing ideas, providing feedback and suggestions to one another, and observing how peers complete tasks (Webb & Mastergeorge, 2003). Students are likely to be proactive about seeking help from their peers if they have a learning goal orientation as opposed to a performance goal orientation (Ames, 1992, 1995; Dweck, 1986; Nicholls, 1984). Students with a learning goal orientation focus on developing understanding and skills. These students are not as concerned with receiving positive feedback as they are with mastering the task. Performance goal oriented students,
however, focus upon the perception of others rather than developing their understanding of the material. These students either want to appear as if they are capable or avoid appearing incapable, regardless of their level of understanding.

In the present study, social dynamics could have played a role in the inability to establish self-efficacy beliefs as a mediating factor between collaborative learning and academic achievement. That is, some of the students who found the material to be overly challenging may have been performance goal oriented. In order to escape judgment from their peers, these students would have appeared as if they understood the concepts, or would have avoided appearing as if they did not understand the concepts, regardless of their true understanding. If this was the case, the performance goal oriented student’s partner would not have sensed the need to offer help during the collaborative learning tasks and, potentially, neither student would have truly engaged in collaborative learning, thus negatively affecting the outcome of the present study.

**Duration of Intervention**

A final factor that affected the ability to establish the mediating effect of students’ topic specific self-efficacy beliefs was the short time period of the intervention. Students who are new to collaborative learning need time to learn how to collaborate (Kelly, 2010), as they do not automatically know how to behave and interact in this new environment (Cohen, 1994a). Recall that the typical mode of instruction for this group of students prior to this instructional unit was lecture and whole class discussion. Therefore, in just three weeks, these students had to adjust to an entirely different mode of instruction, get to know their partner(s), establish a working relationship, and feel comfortable working independently of the instructor. Given more time with these
students, I believe they would have come to a deeper level of comfort working with their peers in a collaborative learning environment, they would have been more at ease working independently of the instructor, and they would have become more engaged in the full collaborative learning process, thus creating an atmosphere of peer modeling and opportunities to experience positive changes in self-efficacy.

**Conclusion**

Several factors affected the ability to establish students’ topic specific self-efficacy beliefs as a mediating factor between collaborative learning and academic achievement for the students who participated in this study. These include the small sample size; the smaller than planned group size which contributed to inconsistent partnership between students, leading to difficulties creating an atmosphere of peer modeling; the social dynamics of collaborative learning groups; and the length of the intervention period. Next I discuss the results of observing the daily class sessions and engaging in subsequent periods of reflection, and how these observations and reflections led to modifications of the collaborative intervention tasks and overall structure of the class.

**Research Questions #3 and #4: Collaborative Task and Class Structure Modifications**

Similar to the methods of design research, I intently observed students and made notes regarding students’ discussions and interactions while they worked in collaboration with their peers during each class period. I also engaged in daily periods of reflection during which I reflected upon the experiences of each class. These observations and reflections served to highlight areas where the collaborative learning tasks could be
improved to help students gain an understanding of the academic material, provide an understanding of the efficacy of the intervention, and allow me to identify instances of peer modeling. Through these daily observations and reflections, I was also able to recognize modifications that needed to be made to the overall class structure and format. Below, I provide recommendations that I believe will strengthen future experiments and discuss why I believe these changes will improve the experience for future instructors and students.

**General Design Recommendations**

When developing future experiments similar to the present study, I recommend that researchers pay close attention to the architectural design of the classroom in which the intervention will occur. Because the students and I had such difficulty with the setup of the classroom in which the present study took place, I now recognize several important characteristics to consider when choosing a classroom in which to implement future interventions. As recommended by Cohen (1994b), I suggest choosing a classroom in which the furniture is not stationary, with tables that can seat four students comfortably, thus allowing students to easily interact with one another in order to foster collaboration. I also recommend that researchers ensure that the classroom provides ample space for the instructor and observers to move about easily without drastically changing the dynamic of student interactions or impeding the collaborative learning process. Because context is such an integral component of the present research study, I strongly suggest that future researchers begin the development of every collaborative learning experiment with consideration of the architectural design of the classroom in which the experiment will occur.
A second suggestion I offer to researchers engaging in qualitative research regarding collaborative learning is to ensure that ample time is provided to implement the intervention and, ideally, instructors should work with their own students. Regardless of who the participating students are, however, I suggest that instructors and researchers implement collaborative learning interventions over the course of one full semester. Both Cohen (1994b) and Kelly (2010) found that it can take students a long time to truly become comfortable working with their peers in a collaborative learning environment. Spending an entire semester with participants would allow the students and instructor the time necessary to develop an atmosphere of trust, respect, comfort and familiarity that would provide a better experience for both the students and the instructor. If everyone in the classroom were given the time necessary to establish familiarity and ease in working with one another, and with the process of collaborative learning, I believe the students would fully engage in the collaborative learning experience. This heightened level of collaboration would create opportunities for students to experience peer modeling, thus positively affecting students’ self-efficacy beliefs. When developing future experiments of this type, researchers should consider both the length of time over which they plan to implement an intervention and the group of students who will participate in the research study.

One final, general recommendation I offer to future researchers is to teach students how to effectively collaborate. Cohen (1994b) warned that students new to collaborative learning are unlikely to possess the skills necessary to collaborate effectively. I strongly suggest that scholars consider allowing plenty of time for this crucial step in the collaborative learning process. Depending upon the group of students
participating in the study, it may be necessary to spend a good deal of time discussing the ideas of collaboration and collaborative learning in the classroom prior to implementing the collaborative learning intervention. I believe that the students who participated in the present study would have benefitted from having more time to work together with their peers on tasks that were not directly related to the instructional unit on statistical sampling, but were intended to help them become more effective collaborators. That is, if these student were allowed time to become comfortable with their peers, to establish relationships while working on tasks that had no bearing on their understanding of the academic material, and to reflect on their experiences and questions about collaborative learning, I believe these students would have been more fully engaged in the collaborative learning process during the instructional unit. This full level of engagement in collaborative learning would have led to experiences that likely would have had a positive effect upon students’ self-efficacy beliefs. In future phases of this design experiment and other experiments, I will rely upon the students’ interactions to help determine the amount of time necessary for students to become comfortable working in this learning environment and, if available, I will provide ample time for this process prior to intervention implementation. As noted by Kalsbeek (Fecso, et al., 1996), the amount of time required will depend upon the context in which the intervention is implemented and will vary for different groups of students.

**Adjustments to Collaborative Tasks and Class Structure**

Beyond the general suggestions I offered for future experiments, I also recommend several specific modifications to the collaborative learning tasks and to the overall structure of the class. Recall that, on the first day of the present experiment, I
realized that some of the wording and phrases used on the collaborative tasks worksheet were not clear to the students. Prior to the first day of the instructional unit, I had three experts in the field of statistics education review all tasks that would be implemented during the three week period. However I did not have anyone review the actual task worksheets, directions, or information that would be presented to the students. In the future, I recommend that instructors and researchers seek the assistance of experts in the field, as well as persons outside the field of study, to review collaborative worksheets for clarity and appropriateness.

Another specific change I recommend for future experiments is to provide explicit, written instructions to students for all collaborative learning tasks. When the students participating in the present study were asked to discuss previous material without written instructions, they experienced great difficulty engaging and collaborating with their partner to complete the task. However, when provided written instructions, directions, or problems, the students seemed more at ease with the collaborative learning process. Cohen (1994b), Gelman and Nolan (2002), and Kelly (2010) believed that these written instructions are useful and serve as a guide for students working in collaborative groups. Therefore, in the future, I suggest providing some form of written instructions, preferably a paper handout, for students to reference as they work together in collaborative learning groups. Note, however, that not all groups of students may experience this same challenge. Therefore, I recommend that instructors and researchers diligently observe students’ interactions to best understand how they receive the use of written instructions and amend accordingly.
Prior to implementing this collaborative learning intervention, I made the mistake of assuming that all students would enjoy working together with their peers as opposed to listening to a lecture. I quickly realized that this was not the case for all students. Some students needed more time than others to become comfortable with this style of learning, while a few students never engaged with their peers at all. This revelation prompted me to reflect upon how I had structured the class, using small group collaborative learning as the primary means of instruction. For future experiments involving collaborative learning, I recommend that instructors and researchers remain open to amending the amount of time students spend working together in small, collaborative groups, while remaining true to the intentions of the research study.

As each group of students is different, I suggest that researchers rely on students’ interactions to dictate how long to allow them to struggle with new material before intervening in some manner. As Edelson (2001) suggested, instructors should intervene on a ‘just-in-time’ basis, providing as little help and guidance as students need. Ideally, I would recommend providing more time than I did in the present study to allow students to work in small groups, think about the task, discuss with their peers, experience cognitive disequilibrium, and create their own ideas around how to best complete the task. This, however, will depend upon the amount of time afforded to instructors to cover the academic material, and how quickly the students come to understand the academic topics. Because the majority of students in the present study were highly receptive to large group, whole class collaboration, I suggest introducing this as a method of collaborative learning as frequently as deemed necessary. Next, I elaborate upon large
group collaboration and offer suggestions for ways to implement this alternative style of collaborative learning.

**Large group collaboration: A form of guided discovery learning.** Large group collaboration combines the characteristics of lecture, whole class discussion, and small group, student led collaborative learning. This type of learning environment is a member of the group of instructional methods called Guided Discovery Learning (Edelson, 2001; Hmelo-Silver et al., 2007; Kuhn, 2007). Other types of Guided Discovery Learning include Problem-Based Learning (PBL) and Inquiry Learning. Hmelo-Silver and colleagues (2007) posited that guided discovery learning is an efficacious means of helping students create deep understanding of various topics and promoting academic achievement. They claimed that unguided experiential learning has not been shown to effectively promote student learning, but that various Guided Discovery Learning approaches structured such that instructors scaffold student learning through considerable guidance can promote high academic achievement. The types of Guided Discovery Learning to which Hmelo-Silver and colleagues referred are heavily collaborative and task-based.

Hmelo-Silver and colleagues (2007) stated that discovery learning can include periods of direct instruction and that, “the teacher plays a key role in facilitating the learning process and may provide content knowledge on a just-in-time basis” (p. 100). Edelson (2001) suggested that a short lecture can be used to convey information once students have had time to struggle with the material and understand why and how the information is pertinent to solving the problem and completing the academic task. Hmelo-Silver and colleagues (2007) also noted that Guided Discovery Learning often
allows students to participate in complex activities and complete tasks they may be otherwise unable to complete on their own.

Recall the experience of the present study when the students in the treatment class began the multistage sampling activity. As the students attempted to work through the task in dyads, I recognized that they were struggling with their partner and initially offered little to no guidance. When we came together as a group, however, we engaged in what would become our first experience with large group collaboration. The students were willing to offer ideas, feedback, varying perspectives, and to collaborate with their peers more so than in past whole class discussions. Students presented their ideas even if they were unsure of the accuracy or how they would be received. One student began by saying, “I don’t know if this is right, but…” and another began with, “I don’t think this is really right….” During this large group collaborative session, the students acted as a united group, working together to help each other understand and offering feedback, thoughts, and ideas for their peers. Eventually, as a large group, the students began to exhibit understanding of the concept of multistage sampling plans. Aligning with the suggestion from Edelson (2001), the students in the treatment class experienced the positive effects of Guided Discovery Learning, or large group collaboration, only after they struggled with their partner during a small group collaborative learning session.

This experience of students taking over the class and responsibility for helping their peers understand the material relates to the idea of mathematical authority (Hamm & Perry, 2002). Many students, at all levels of education, do not recognize that mathematics, and statistics, are disciplines constructed by humans (Schoenfeld, 1992; Stodolsky, 1988). However, during this large group collaborative session, it was the
students who had the statistical authority to create and develop their own knowledge of the concept of multistage sampling. This type of classroom environment, in which students take over and begin to teach and learn from one another, is the type of environment instructors at all levels of education would often like to achieve. It was exciting to experience this with the treatment group of students and to recognize that large group collaboration is an effective alternative to small group collaborative learning.

Hmelo-Silver and colleagues (2007) offered suggestions to instructors for ways to scaffold student understanding within a Guided Discovery Learning environment. They recommended that instructors act as a model or coach for students, using scaffolding and, over time, reducing the amount of support provided to students. They noted that instructors employing these techniques often asked students to explain their ideas, a strategy that helped students talk through their thoughts and identify where they were struggling to understand certain concepts. The authors also recommended that instructors use various question prompts to incite discussion among students in the class and to identify where he or she should offer just-in-time information as necessary. Hmelo-Silver and colleagues suggested that tracking problem-solving processes on a whiteboard in the classroom can help to reduce the cognitive load of students, thus allowing students to focus upon the activity itself and to complete the tasks necessary to meet the academic goals. Tracking this information for students to reference is also a way of fostering collaborative discussion, as it serves as a visual reminder of the various aspects associated with completing the task.

The overwhelming positive student response to large group collaboration was encouraging for me as an instructor. When I recognized that not all students enjoyed
working with their peers in small groups, I was concerned about how I could provide most, if not all, students with a productive learning experience. As a result of students’ struggles with small group collaborative learning, the students and I discovered large group collaboration, a method of instruction I plan to use in the future. I recommend that researchers studying collaborative learning consider using the guided discovery method of large group collaboration as often as student interactions deem necessary.

**An Ideal Collaborative Learning Intervention Study**

My overall goals with the present study was to learn a great deal about how to conduct a collaborative learning intervention research study, to consider what I would do differently to improve the study for future iterations, and to modify the study in ways that allow for those improvements. When I have the opportunity to conduct a similar intervention study in the future, there are several aspects I would change based upon the qualitative and quantitative data results and my experiences of the present study. I have mentioned several suggestions for improvement of the present study. I will include those, as well as other aspects, in presenting my ideal collaborative learning intervention study.

One aspect that would optimize an intervention study of this type would be to implement the study in a classroom with large, moveable tables and chairs that comfortably seat four students. Ideally, the tables would be round so that students can sit in a circle and face one another, a factor Cohen (1994b) believed to be conducive to collaborative learning. The classroom itself would be large enough to accommodate 24 to 28 students, or 6 to 7 groups of four. The classroom would also have ample space between the tables and chairs to allow the groups to work without disruption from one another, as these disruptions can impede collaboration (Cohen, 1994b). It would be large
enough so that the instructor and any other observers could move freely about the room without changing the dynamic of the small group collaborative learning sessions.

Implementing the study in a suitable sized classroom with an architectural layout amenable to collaborative learning would allow groups of four students to work together in a comfortable environment. When students are assigned to groups of four, the instructor can assign each student a specific role to play during the collaborative learning exercise, a recommendation Cohen (1994b) believed helps foster collaboration. In the present study, I was not able to arrange students in groups of four. Instead, the architectural layout of the classroom forced me to pair students in dyads. This arrangement made it impossible to assign students specific roles such as facilitator, reporter, materials manager, etc. In an ideal collaborative learning intervention, each group would consist of four students and I would assign each student a specific role to play, thus promoting a collaborative learning environment.

Ideally, I would work with a group of students over the period of a full semester to implement the collaborative learning intervention. This would allow students the time needed to become comfortable with me and with this potentially new learning environment (Kelly, 2010). I would also be the instructor of record for the class in which I implemented the intervention. I believe this instructor-student relationship would encourage a large number of students to participate in the study and would help students begin to feel at ease learning in this environment sooner than if I was not the primary instructor. As the instructor for the class, I would have the authority to offer some type of reasonable incentive to encourage ongoing student participation. I believe that
offering an incentive for students to consent to participate and to remain in the study for the full research period would greatly reduce participant attrition.

Another aspect that I believe would make for an ideal study would be to have one or two observers, besides the instructor, attend each class period. The job of these additional personnel would be to observe the students as they worked in small, collaborative groups and to note their interactions and verbalizations. They would specifically note the appropriate use of academic language during student discussions, the occurrence of negotiations among the students in each group, the presence of productive argumentation among group members, and times when positive social interdependence was present, all indicators of collaborative learning (Chinn, 2006; Cohen, 1994b; Dillenbourg, 1999; Erkens et al., 2006; Grover, 2010; Johnson & Johnson, 1989; Ploetzner et al., 1999). I believe that having additional personnel in the classroom to note when these phenomena occur would allow me, the instructor, to focus upon ensuring that the collaborative learning tasks and format of the class help create an atmosphere of collaboration for the students.

Each post-class period of reflection would begin with a discussion between myself and the additional observers to debrief regarding the events of the day. This would serve as our time to collaborate and determine how to best modify the tasks and class format for subsequent class periods. After my discussion with the observers, I would engage in a period of personal reflection during which I would review the written observation notes from the class and would note my own experiences, thoughts, and ideas for intervention improvement. I would then make the final collaborative learning task and class format changes for the next class period, basing these modifications upon the
written observations of all observers, the discussion among myself and the additional observers, and my personal reflections.

I believe that the effect of the intervention may be best assessed at a combination of the group and individual levels. Because this intervention is intended to investigate the relationship between collaborative learning and self-efficacy beliefs, I feel that assessing students both individually and collaboratively could shed greater light on how these two phenomena are related. Perhaps students who have been learning in a collaborative environment would perform better on this combination of assessments than on an individual assessment alone. Therefore, in an ideal study, I would administer the pre and post knowledge assessments to the treatment group of students both individually and as a collaborative group exercise. That is, I would have each student within each group complete the pre and post assessments on their own and, then, I would give each group one common assessment to be completed collaboratively. The group assessment would cover the same concepts as the individual assessment, but would not include exactly the same assessment items. The individual assessment would focus upon basic, declarative knowledge and the group assessment would include open-ended questions that would require students to exhibit more complex procedural knowledge. Use of both individual and group assessments is common practice in the K-12 setting (see Billington, 1994) and is becoming more prevalent in the college classroom (see Artino & Ioannou, 2010; Kelly, Baxter, & Anderson, 2010). Each student’s pre and post assessment scores would be a composite of his or her individual assessment score and the group assessment score.
As recommended by Cohen (1994b), in order to incite discussion and collaboration among students, each class period I gave every dyad only one copy of the worksheet with which to complete the collaborative tasks. My intention was to prompt students to engage with one another instead of working individually on their own paper. However, providing only one worksheet could have a negative impact upon students’ opportunities to review the worksheet outside of class. Therefore, in an ideal study, I would encourage students to make group notes on the worksheet and, if necessary, take their own personal notes during the collaborative learning sessions. I would then take time at the end of each class period to make photocopies of the all groups’ worksheets and individual notes and provide each student within each group a copy of the work that his or her group completed. I believe it is important to provide students with copies of their work, as opposed to a worksheet completed by the instructor, so that they can recall their thought processes, shared ideas, and perspectives provided by other members of the group.

In order to truly understand the students’ experiences and to gain an understanding of what they learned and how their self-efficacy beliefs were influenced by the intervention, I would take the time to interview a small subset of students after completion of the collaborative learning intervention. I believe the information gathered during these informal interviews regarding students’ experiences during the collaborative learning exercises could fill some gaps in my understanding of the interplay between collaborative learning and self-efficacy beliefs. I also believe that these data could help to inform how subsequent iterations of this type of intervention study should be implemented, offering ideas for improvement from the student perspective. During these
interviews, I would review with the students some of their work from the collaborative learning sessions. This would allow them to clarify any misunderstandings I may have and to elaborate on their thought processes and experiences. Ideally, I would conduct these interviews in a focus group format, talking with students from the same collaborative group who worked together during the semester.

Finally, the ideal intervention study would include a balance of instructional types. This balance would change for each implementation and each group of students, as this balance is very dependent upon the needs of the students (Fecso et al., 1996). Finding and implementing just the right amount of small group collaborative learning, mixed with whole class discussion, mini-lecture, and large group collaboration would provide for students the type of learning environment necessary to gain a deep understanding of the material, could foster a classroom atmosphere of collaboration (Edelson, 2001; Hmelo-Silver et al., 2007), and would keep students more interested than lecture alone (Kelly, 2010).

**Ideal task design.** The process of designing collaborative learning tasks requires much forethought and planning (Cohen, 1994b). Cohen recommended several aspects instructors should consider when developing collaborative learning tasks. She noted that the type of tasks and how they are developed depends on several criteria, including:

1. What does the instructor want students to learn? For example, if the intention is that students understand the relationship between a sample and a population, the tasks should be designed such that students can discover and discuss the nature of this relationship. If the goal of the task is to encourage social engagement among
peers, the instructor should develop a highly engaging task that will encourage students to interact with one another.

2. What tasks are best implemented in a collaborative learning environment? Cohen cautioned that not all tasks are collaborative tasks. Instructors must determine which tasks are best suited to be implemented in a collaborative learning setting.

3. What level of students will work to complete the task? Cohen cautioned that, when working with younger students, the instructor may have to provide more guidance and instruction throughout the task. However, college level students should be able to proceed through the task with less intervention from the instructor.

4. How complex is the concept? For more complex concepts, Cohen recommended that students be provided with a series of tasks and activities, as opposed to a single task, to ensure that they fully understand the central idea and can apply their knowledge and understanding to other scenarios.

5. In what context will the task be implemented? That is, instructors should consider how many students are in the class, how the students will be grouped, how large each group will be, etc. These factors will affect the development of the task.

6. What role will the students within each collaborative group fulfill? Instructors should consider which roles the collaborative task requires and how to assign those roles. For example, does the task require a facilitator, reporter, and checker? Is a materials manager also required? These aspects, taken into consideration along with the context and group of students, will impact how the task should be designed and implemented.
In an ideal study, I would collaborate with experts in the fields of statistics education and collaborative learning to ensure that the tasks I develop are designed to help students learn the statistical concepts, as well as foster productive collaboration among students. I would begin the task design process by writing a goal statement that includes all key concepts I want the students to learn. Once these goals are set, I would enlist the help of collaborative learning experts to determine which concepts are best suited for a collaborative learning environment. As Cohen (1994b) mentioned, not all tasks are collaborative tasks. I would need to determine which concepts require multiple perspectives and develop collaborative learning tasks that are directly connected to those key concepts.

Because I would be working with college students, I would develop tasks that require little to no intervention from the instructor. However, during implementation, I would be keenly aware of how the students were receiving the tasks and would intervene only as needed. For complex concepts, I would follow Cohen’s (1994b) advice and create a series of tasks to help students develop a full understanding of the concept. These sets of tasks would be designed such that students could transfer their knowledge and understanding to similar future tasks.

Once I knew the class size and group of students with which I would be working, I would make sure that the tasks are developed accordingly. For example, if I have a class of more than 30 students, I would develop tasks that would work well for groups of five, or maybe even six, students. Although Cohen (1994b) warns against assigning students to groups larger than five, when working with a very large class it may be warranted to increase the group size slightly. However, when working with a class of
less than 30 students, I would develop tasks that were suited for groups of four students. The intended group size would affect the way in which the task is developed and the various roles that are necessary for each task. It is important to consider the roles required for each task during the task development stage so that each student within the collaborative group understands their place in the group as well as their responsibilities.

Once I had completed the task design process, I would have experts in the fields of statistics education and collaborative learning review all collaborative learning tasks. This would ensure that the tasks are designed to help students learn the key concepts and to encourage collaboration among students. This review by experts is crucial to establishing and helping to increase the efficacy of the collaborative learning treatment.

Conclusion

I have discussed several general and specific recommendations for future experiments of this type and have presented what I believe to be an ideal collaborative learning intervention study. I believe these suggestions help to strengthen this piece of research and provide future researchers a strong base from which to develop their own experiments. Next, I discuss modifications made to the foundational learning theory, changes that are based upon the observations and reflections from the present study.

**Research Question #4: Modifications to Foundational Learning Theory**

Recall that the original foundational learning theory stated:

1. With the treatment group of students, I would begin by discussing what the term collaboration meant, how collaborative groups could be effective and productive, and what was expected from each member of a collaborative group.
2. Daily, I would use collaborative learning as an instructional method to allow students to work, discuss, and collaborate with their peers. I believed this type of learning environment would give students the freedom to express their thoughts and ideas, feel safe to be wrong, and generate their own comprehension of statistical topics. I also believed that this opportunity to collaborate with their peers would help students feel a greater sense of self-efficacy regarding their statistical knowledge and understanding of the concepts associated with statistical sampling, as well as their ability to reason about statistical ideas. This collaborative learning time would serve as the primary means of instruction for the students, with little to no intervention from me, the instructor.

3. I would allow a short period of whole class discussion at the end of each class. The intent of this was to ensure that the students understood the material as I expected, to share insights the students had gained while working collaboratively with their peers, and to answer any outstanding questions regarding the statistical topic of the day.

Given the recommendations discussed above, I modified the theory as follows:

1. Prior to designing a collaborative group intervention, the instructor should consider the architecture of the classroom in which the intervention will take place. The ideal classroom would include movable tables at which four students could sit comfortably and move rather freely. The classroom would also include moveable chairs and ample space for the instructor to move about the classroom and observe students working without impeding student work and interactions.
2. The instructor should not assume that students know how to work together collaboratively, or that all students want to work with their peers in a small group setting. Prior to engaging in tasks related to the academic topic, a whole class discussion of the meaning of collaboration, how collaborative groups can be effective and productive, and what is expected from each member of a collaborative group should occur. This discussion should include how collaborating with student peers can be beneficial to the learning experience. The instructor should allow students time to express their concerns and apprehensions in regard to working together with their peers, and he or she should try to help alleviate these concerns.

3. On a daily basis, the instructor should use collaborative learning as an instructional method to allow students to work, discuss, learn from, and collaborate with their peers. However, the instructor should be very aware of how the students are accepting this type of learning. Once the instructor has allowed students ample time to struggle with the material and/or this mode of instruction, the instructor may need to intervene with whole class discussion or large group (i.e., whole class) collaboration. During large group collaboration, the instructor should intervene as little as possible. Achieving this balance between allowing students the time to experience sufficient cognitive disequilibrium or discomfort, and helping students attain maximum understanding may vary based upon the context in which the intervention is implemented.

4. During collaborative tasks, it is important that students have some form of written, explicit instructions to follow or questions to answer. Without these
written instructions, students may struggle to follow directions and to successfully complete the tasks at hand. Students may also be less inclined to engage with their peers if they do not have a set of written instructions to follow.

5. At the end of each class period, the instructor should bring the class together for a final whole class discussion to ensure that students understand the material as expected, to allow student to share insights they gained while working collaboratively with their peers, and to answer any outstanding questions students may have regarding the topic of the day.

The methods of qualitative research allowed me to make informed modifications stated above. I believe this revised learning theory provides a stronger foundation for the next phase of this and other experiments of this type. Beyond my own use, it is my hope that researchers who wish to implement a similar experiment in their classrooms will use this new learning theory as a basis upon which to develop an experiment that best suits the needs of their students. I believe the detail and explicit language of the modified learning theory is a vast improvement from the original version.

**Limitations and Future Directions**

In the present study, I attempted to establish that college statistics students working in small, collaborative, student led, peer groups in the classroom would show a statistically significant increase in topic specific self-efficacy beliefs. I also sought to demonstrate that students’ topic specific self-efficacy beliefs mediated the relationship between collaborative learning and academic achievement of the students studying statistical sampling. Unfortunately, due to several factors discussed above, I was unable to provide support for either of these hypotheses. In this section I discuss various
limitations of the present study, as well as suggestions for future work to be conducted to study the phenomena of self-efficacy beliefs, collaborative learning, and academic achievement in concert with one another.

The present study was subject to various potential threats to internal validity. These included attrition of participants, selection bias, and repeated testing. While none of these seemed to greatly affect the internal validity of the study, I offer means of addressing these in future studies such that they do not adversely affect the research study outcomes.

The attrition of participants did not affect the balance of students between the control and treatment groups who fully completed the study. It is possible, although unlikely, that students who did not complete the post unit survey and assessment were in some way systematically different from students who did complete these measures. Participant dropout did, however, affect the study results in that I had a very small sample from which to draw conclusions. In future studies, I recommend frequently encouraging students to remain active in the study and considering the use of a non-academic incentive to promote students continued participation throughout the entire research period.

Educational researchers often conduct quasi-experimental research studies within classrooms as a means of overcoming various obstacles that school settings present. One of these obstacles is that students are pre-assigned to classrooms, thus not allowing for random selection of participants to groups. However, not randomly assigning students to treatment and control groups can lead to multiple problems, including selection bias. The only way to ensure that no factor other than the intervention itself influences outcome
differences between groups is to conduct a true experiment in which researchers can randomly assign participants to groups. Because it is virtually impossible to conduct a true experiment within the confines of typical school environments, researchers conducting quasi-experiments attempt to choose treatment and control groups such that they are as equal as possible across all characteristics. As a quasi-experimental researcher, I believe this is the best method of attempting to eliminate the internal validity threat of selection bias.

When a research study includes several testing instances over the course of the study, participants may become familiar with the test itself or may simply become conditioned to being tested. These issues associated with repeated testing can affect test scores in ways that may result in outcomes that are not necessarily due to the intervention itself. In the present study, I was aware of this potential threat to internal validity and at no time did I reveal to the students the correct knowledge assessment answers. In future research studies of this type, I recommend that instructors and researchers do not provide test answers to participants until after the end of the research study. I also recommend testing participants as infrequently as possible to reduce the threats associated with conditioning.

As with many research studies, these results can be applied to the participating students and extrapolated only to students with similar demographics studying within a context very much like that of the present study. One of the major limitations of this study is that the small sample size that made it challenging to draw any significant conclusions from the quantitative data. In order to extrapolate the results of future
studies to other students, a large scale research study should be conducted based upon the revised learning theory and modified collaborative learning tasks.

The sample of participating students in the present study was not only small; it also lacked racial diversity. Recall that 95% of the participating students in these two classes self-reported as white. At the university where I conducted the present study, 88% of undergraduate students self-reported as white. Thus, the racial composition of the sample for the present study was similar to that of the undergraduate student body. However, I believe the present study is limited by the fact that the sample was comprised of participants from primarily one race. Therefore, these research findings cannot be extrapolated to a mixed-race sample. In future studies, it is crucial to recruit students of various races and ethnicities so that the results can be extrapolated to a more diverse sample and to allow investigation of the differences in the relationships among self-efficacy beliefs, collaborative learning, and academic achievement in statistical sampling among these ethnic groups. In order to accomplish this, future studies should be conducted on college campuses where the racial composition of students is more diverse than was the case for the present study.

Another limitation of the present study which contributed to the inability to establish statistical significance for the relationships between collaborative learning and students’ topic specific self-efficacy beliefs is the short time period I was able to spend with the participating students. In my past experiences as a university instructor, I have used collaborative learning as the primary method of instruction in various undergraduate statistics courses. Through these experiences, I have come to find that it can take students several class meetings working in small groups with their peers to become
comfortable with this often new style of learning. Working with the students in the present study for only three weeks, or six class periods, hindered the chance of establishing statistically significant relationships among the phenomena under investigation. The students did not have the time required to establish relationships with one another and fully engage in the collaborative learning process.

For future studies, I suggest two possible solutions. Ideally, the instructor who implements the intervention should be the primary instructor who teaches the course. That is, he or she should be the person who has been working with the students throughout the semester, who has established a rapport and level of comfort with the students, and whom the students trust and respect. The instructor should use collaborative learning as a method of instruction throughout the semester so that students become very comfortable and familiar with this style of learning. If this is not possible, then I suggest that the instructor who implements the intervention should be present in the classroom with the students for the full semester, possibly assisting the primary instructor and working with the students to establish positive working relationships. I suggest that the assisting instructor implement the intervention over a long time period, preferably a full semester. This longer period of implementation would provide students the time necessary to become comfortable working with their peers in a collaborative learning environment and would allow for an atmosphere of peer modeling, thus increasing the opportunities for students to experience positive changes in self-efficacy.

Finally, I believe that the present study was limited because I served as both the instructor and the sole researcher. As mentioned previously, the instructor of record for this course was asked to be absent during the three weeks of intervention implementation.
This was to preserve the confidentiality of participating students so that the instructor would neither reward nor punish students based upon their participation status. Therefore, she could not serve as a resource to help observe and take notes during the class sessions. I did not have the resources to employ someone to observe in the classroom along with me and to take notes during the experiment. However, after completing the full study, I now recognize the benefits of having at least one other person in the classroom to help gather the observational data during each class period. In future studies, I suggest that a second person be present during all class sessions. This person would not interact with the students, but would serve to make observations and take notes during each class. This person should be someone who has no bearing on students’ grades and has no direct association with the course itself. If a person is not available to serve in this capacity, I recommend videotaping the class periods and reviewing the sessions after each class. Though this would not be as effective as having a second observer in the classroom, it would allow the instructor to review the interactions of the students and to possibly discover student behaviors he or she may not have noticed during classroom observations.

**Future Directions**

Because the results of the present study did not support my hypotheses regarding the relationships among collaborative learning, students’ self-efficacy beliefs, and academic achievement in statistical sampling, future researchers should consider investigating these relationships in additional ways. Perhaps it would be beneficial to specifically investigate the peer modeling aspect of collaborative learning and its effect upon students’ self-efficacy beliefs. That is, future researchers could design additional
pre and post surveys that tap into the phenomenon of peer modeling to determine how it affects the relationship between collaborative learning and the academic achievement of students studying statistical sampling. Peer modeling could be tested as a mediating factor between collaborative learning and self-efficacy beliefs, thus potentially affecting academic achievement.

Another alternative may be to consider the relationship between collaborative learning and self-efficacy beliefs in the opposite direction than in the present study. Various scholars have studied how students’ self-efficacy beliefs affect their performance in small group settings (Ruys et al., 2010; Sins et al., 2008; Wang & Lin, 2007). None of these studies, however, focused specifically on collaborative learning as the small group learning environment. It may be that students with higher self-efficacy beliefs engage in more effective collaboration with their peers than students with lower self-efficacy beliefs, thus leading to better academic performance. Therefore, future researchers may consider the effect of students’ self-efficacy beliefs upon collaboration and how that, in turn, affects the academic achievement of students studying statistical sampling.

In the present study, I was not able to establish that college statistics students working in small, collaborative, student led, peer groups in the classroom statistically significantly increased their topic specific self-efficacy beliefs. Nor was I able to establish students’ topic specific self-efficacy beliefs as the mediating factor between collaborative learning and academic achievement in statistical sampling. However, I have noted several limitations of the present student that I believe impacted the ability to establish these statistically significant relationships. I have also made several suggestions
for future studies that I believe will strengthen the research of the relationships among these three phenomena.
CHAPTER 6

CONCLUSION

Approximately half of undergraduate students who begin a degree program at a 4-year college or university graduate within six years (U. S. Department of Education, 2010). In order to increase this graduation rate, undergraduate instructors and students must develop methods of promoting academic achievement, thus encouraging college degree completion (U. S. Department of Education, 1983). For undergraduate statistics students enrolled in an introductory course, the topics associated with statistical sampling can be especially difficult to understand (Fecso et al., 1996; Gelman & Nolan, 2002; Nguyen, 2005; Utts & Heckard, 2006; Yilmaz, 1996). Students often find these topics to be overly abstract and complex (Nguyen, 2005; Yilmaz, 1996), and instructors have difficulty teaching these concepts to students with varying mathematical and statistical backgrounds (Fecso et al., 1996; Yilmaz, 1996). Two phenomena that have been shown to increase academic performance in various academic domains and for specific topics within those domains are students’ self-efficacy beliefs (Bandura, 1995; Pajares, 2006) and collaborative classroom learning (Johnson & Johnson, 1989).

Based on research in the field of self-efficacy, or, peoples’ beliefs regarding their abilities to be successful in their pursuits (Pajares, 2006), it has been shown that students with higher self-efficacy beliefs for an academic domain, topic, or task tend to achieve at higher levels within that domain or on that topic or task than students with lower levels of self-efficacy beliefs (Bandura, 1997; Pajares, 1996; Pajares & Urdan, 2006). Self-
efficacy scholars have suggested that educators and students understand the ways in which students’ self-efficacy beliefs can be improved in order to bolster these beliefs as a mechanism of creating academic success (Chemers et al., 2001; Pajares, 1996). One of the primary sources of increasing self-efficacy beliefs is peer modeling, or the act of students observing their peers successfully complete a task through effortful perseverance (Bandura, 1995). It has been shown that collaborative learning is a more successful method of promoting academic achievement than individual or competitive learning (Johnson & Johnson, 1989; Johnson et al., 1981). When working effectively, a collaborative learning environment is distinguished by the presence of positive social interdependence, which results in promotive interaction (Johnson & Johnson, 1989). According to social interdependence theory, peer modeling is one of the mechanisms by which learning and achievement occur. Collaborative learning both depends on and fosters promotive interaction, the key characteristic of positive social interdependence. When students work together in small, collaborative groups and have the opportunity to observe their peers working to achieve academically, they often model their successful peers. Because peer modeling is one of the primary sources through which students increase their self-efficacy beliefs (Bandura, 1995), it follows that the peer modeling aspect of collaborative learning can help students increase their self-efficacy beliefs, leading to greater academic success.

Both collaborative learning and increased self-efficacy can help undergraduate statistics students studying statistical sampling reach high academic achievement. In the present study, I hypothesized that the self-efficacy beliefs of undergraduate students’ studying statistical sampling would increase as a result of working with their peers in
small, collaborative groups within the classroom. I also proposed that students’ topic specific self-efficacy beliefs mediated the relationship between collaborative learning and academic achievement for undergraduate statistics students studying statistical sampling. In order to provide support for these hypotheses, I administered pre and post surveys and assessments to gather quantitative data in an effort to answer the following research questions:

1. Does working in small, collaborative, student led, peer groups in the classroom increase the topic specific self-efficacy beliefs of undergraduate level statistics students?

2. Do topic specific student self-efficacy beliefs mediate the relationship between collaborative learning and the academic achievement of undergraduate level statistics students studying statistical sampling?

While I was not able to provide statistical support for these two hypotheses, through the use of qualitative research methods I was able to determine how to best revise the collaborative learning intervention tasks and the foundational learning theory such that they are more efficacious, impactful to student learning, can help increase students’ self-efficacy beliefs, and applicable to the context in which the intervention is implemented. This methodology helped me answer the final two research questions:

3. How should the collaborative tasks intended to help undergraduate students learn topics related to statistical sampling be redesigned and implemented such that they foster the topic specific self-efficacy beliefs and, thereby, promote higher academic achievement of undergraduate students?
4. How should the overall research design be revised for use in future experiments of the same type implemented in similar contexts?

Not only was I able to modify the collaborative learning tasks after each class period and implement these changes during subsequent class sessions of the present study, I also used the qualitative data gathered during the full six day experiment to inform how to best revise the statistical sampling collaborative learning tasks and the learning theory for use in future iterations of this experiment and future experiments similar to the present study.

I believe the results of the present study contribute to the contemporary literature regarding self-efficacy beliefs and collaborative learning in several ways. First, the qualitative results provide a starting point from which teachers and researchers can develop future experiments to study these two phenomena together in their classrooms. Second, this study highlights certain criteria and potential pitfalls that require special attention when developing and implementing a study of this type. Finally, because there is little empirical research examining the relationship between students’ self-efficacy beliefs and collaborative learning, the present study encourages scholars to consider future research that investigates these two phenomena in concert with one another.
Appendix A

Student Recruitment Script

“Hello, my name is Jane Robertson and I am a graduate student at the University of North Carolina at Chapel Hill. I am conducting a research study about the relationship between college students’ beliefs about statistics and collaborative learning. I am here to ask whether you would be willing to participate in this research study.

Participation would involve three things:

1. your reading and signing a consent form that provides information about the research study and your rights as a participant,
2. taking a short survey now and a short survey at the end of the semester,
3. taking an assessment about your prior statistical knowledge of the unit we will cover, and
4. agreeing to allow your instructor to release your unit exam grade and final course grade to me after he or she has submitted your grade to the Office of the University Registrar.

I will be giving you copy of the consent form to read and sign, and a second copy of the consent form for you to keep for your records.

I will never use your name in any kind of publication or presentation. In addition, your participation will be confidential—your instructor will not know who is and who is not participating until after he or she has submitted your final course grade to the Office of the University Registrar.

If you decline to participate, this will have no impact upon your standing in this class.

If you begin the research study now and choose to leave the research study between now and the end of the semester, there is no penalty for doing so.

I will distribute a set of materials to everyone. The packet contains the two copies of the consent form, the survey, and the prior knowledge assessment. You can look at the survey and assessment to see what they involve before you decide about participation. Once you have read the consent form, if you are interested in participating, please indicate your willingness to participate and sign the form. Then please complete the short survey and assessment. If you are not interested in participating after reading the consent form, please indicate that you wish not to participate and sign the consent form. Once you are finished, please wait for your peers to finish. I will collect all material packets at the same time.

Thank you for your willingness to consider being in my research study.”
Appendix B

Description of Study and Student Consent Form

University of North Carolina at Chapel Hill
Consent to Participate in a Research Study
Adult Participants
Consent Form Version Date: 8/27/11
IRB Study # 11-1700
Title of Study: Self-efficacy and Collaborative Learning: An Intervention Study
Principal Investigator: Jane Robertson
Principal Investigator Department: School of Education
Principal Investigator Phone number: 919-357-1670
Principal Investigator Email Address: janer009@email.unc.edu

What are some general things you should know about research studies?
You are being asked to take part in a research study. To join the study is voluntary. You may refuse to join, or you may withdraw your consent to be in the study, for any reason, without penalty.

Research studies are designed to obtain new knowledge. This new information may help people in the future. You may not receive any direct benefit from being in the research study. There also may be risks to being in research studies.

Details about this study are discussed below. It is important that you understand this information so that you can make an informed choice about being in this research study.

You will be given a copy of this consent form. You should ask the researchers named above, or staff members who may assist them, any questions you have about this study at any time.

What is the purpose of this study?
The purpose of this research study is to learn about college students’ beliefs about and achievement in the domain of statistics.

You are being asked to be in the study because you are enrolled in an undergraduate Basic Statistics course.

How many people will take part in this study?
A total of approximately 60 people at Appalachian State University will take part in this study.

How long will your part in this study last?
Your participation in the study will last approximately six 50-minute class periods during the Fall 2011 semester.
What will happen if you take part in the study?
If you choose to participate in this study, you will be asked to complete a VERY short survey and a pre knowledge assessment today. Over the next 3 weeks, you will participate in class as normal.

After 3 weeks, or 6 class periods, you will be asked to complete a VERY short survey and a post knowledge assessment.

What are the possible benefits from being in this study?
Research is designed to benefit society by gaining new knowledge. You will not benefit personally from being in this research study.

What are the possible risks or discomforts involved from being in this study?
The only risk to you is that of a breach of confidentiality. There are several measures in place to minimize this risk. Student consent forms, surveys, and assessments will be stored in a locked file cabinet until the end of the study, at which time they will be shredded and disposed of. Once all data have been collected in paper form, data will be transferred to an electronic database and no identifying information will be included in the database.

The only possible discomfort would be a feeling of subtle pressure to participate if you and your fellow students thought that your instructor would know who did, and who did not, choose to participate, and that information might affect your instructor’s perceptions of the students, and perhaps their grades. To minimize this possible discomfort, steps will be taken to ensure that your instructor will not know who is participating in the study. Your professor will not be in class during any part of the 3 week study period. She will never know whether or not you participated in the study.

There may be uncommon or previously unknown risks. You should report any problems to the researcher.

What if we learn about new findings or information during the study?
You will be given any new information gained during the course of the study that might affect your willingness to continue your participation.

How will your privacy be protected?
The Principal Investigator is responsible to protect your privacy and the confidentiality of your information. So that your decision to participate or not remains private, all students will receive a packet of materials to review. Each student has the option to complete the paperwork and surveys or not. As noted above, once identifiable data have been collected, it will remain in a locked filing cabinet for the duration of the research study. Only the Principal Investigator will have access to identifiable data.

Participants will not be identified in any report or publication about this study. Although every effort will be made to keep research records private, there may be times when federal or state law requires the disclosure of such records, including personal
information. This is very unlikely, but if disclosure is ever required, UNC-Chapel Hill will take steps allowable by law to protect the privacy of personal information. In some cases, your information in this research study could be reviewed by representatives of the University, research sponsors, or government agencies (for example, the FDA) for purposes such as quality control or safety.

**What if you want to stop before your part in the study is complete?**
You can withdraw from this study at any time, without penalty. The investigators also have the right to stop your participation at any time. This could be because you have failed to follow instructions or because the entire study has been stopped.

**Will you receive anything for being in this study?**
You will not receive anything for taking part in this study.

**Will it cost you anything to be in this study?**
It will not cost you anything to be in this study.

**What if you have questions about this study?**
You have the right to ask, and have answered, any questions you may have about this research. If you have questions about the study (including payments), complaints, concerns, you should contact the researchers listed on the first page of this form.

**What if you have questions about your rights as a research participant?**
All research on human volunteers is reviewed by a committee that works to protect your rights and welfare. If you have questions or concerns about your rights as a research subject, or if you would like to obtain information or offer input, you may contact the Institutional Review Board at 919-966-3113 or by email to IRB_subjects@unc.edu.
Participant’s Agreement:
I have read the information provided above. I have asked all the questions I have at this time.

Please check one of the following:

___ I agree to participate in this study.

OR

___ I decline to participate in this study.

______________________________________________________
Signature of Research Participant

______________________________________________________
Date

______________________________________________________
Printed Name of Research Participant

______________________________________________________
Signature of Research Team Member Obtaining Consent

______________________________________________________
Date

______________________________________________________
Printed Name of Research Team Member Obtaining Consent
Appendix C

Student Self-efficacy Beliefs Survey – Pre-Instruction

1. I believe I will receive an excellent grade on the project related to statistical sampling.

   A. Strongly agree
   B. Agree
   C. Somewhat agree
   D. Neutral
   E. Somewhat disagree
   F. Disagree
   G. Strongly disagree

2. I’m certain I can fully understand the most difficult material regarding statistical sampling that was presented in the readings.

   A. Strongly agree
   B. Agree
   C. Somewhat agree
   D. Neutral
   E. Somewhat disagree
   F. Disagree
   G. Strongly disagree

3. I’m confident I can understand the basic concepts regarding statistical sampling.

   A. Strongly agree
   B. Agree
   C. Somewhat agree
   D. Neutral
   E. Somewhat disagree
   F. Disagree
   G. Strongly disagree
4. I’m confident I can understand the most complex material regarding statistical sampling that was presented by the instructor.

A. Strongly agree
B. Agree
C. Somewhat agree
D. Neutral
E. Somewhat disagree
F. Disagree
G. Strongly disagree

5. I’m confident I can do an excellent job on the tasks related to statistical sampling.

A. Strongly agree
B. Agree
C. Somewhat agree
D. Neutral
E. Somewhat disagree
F. Disagree
G. Strongly disagree

6. I expect to do well on this unit about statistical sampling.

A. Strongly agree
B. Agree
C. Somewhat agree
D. Neutral
E. Somewhat disagree
F. Disagree
G. Strongly disagree
7. I’m certain I can master the skills being taught in this unit about statistical sampling.

A. Strongly agree
B. Agree
C. Somewhat agree
D. Neutral
E. Somewhat disagree
F. Disagree
G. Strongly disagree
Appendix D

Pre and Post Statistical Knowledge Assessment

1. The state of North Carolina surveyed 1000 parents of children attending a daycare facility within North Carolina to determine whether the state government should increase funding to daycare providers throughout the state.

   a. What is the population of interest in this survey?
      i. The 1000 surveyed parents of children attending a daycare facility within North Carolina
      ii. The children of the parents surveyed
      iii. All parents of children attending a daycare facility within North Carolina
      iv. All children attending a daycare facility within North Carolina

   b. What is the sample in this survey?
      i. The 1000 surveyed parents of children attending a daycare facility within North Carolina
      ii. The children of the parents surveyed
      iii. All parents of children attending a daycare facility within North Carolina
      iv. All children attending a daycare facility within North Carolina

2. Which type of bias is defined in each of the following scenarios?

   a. Participants respond differently from how they truly feel.
      i. Response bias
      ii. Nonresponse bias
      iii. Selection bias

   b. The method for selecting the participants produces a sample that does not represent the population of interest.
      i. Response bias
      ii. Nonresponse bias
      iii. Selection bias
3. You are asked to conduct a survey to determine what percentage of students at Appalachian support the legalization of marijuana. You must survey enough students such that the conservative margin of error is 5% or lower. What is the minimum number of students you must survey (i.e., minimum sample size)?
   i. 100
   ii. 400
   iii. 20
   iv. 1000

4. You decide to survey 500 Appalachian students and find that 79% support the legalization of marijuana. Using the conservative calculation for the margin of error, what is the approximate 95% confidence interval for the percentage of all Appalachian students who support the legalization of marijuana?
   i. 70.25% to 88.10%
   ii. 77.32% to 81.89%
   iii. 79.54% to 85.22%
   iv. 74.53% to 83.47%

5. In each of the following scenarios, indicate the type of sampling used:
   a. Two farmers from each state in the United States are randomly chosen to serve on the National Farmers Coalition board.
      i. Simple random sampling
      ii. Stratified random sampling
      iii. Cluster sampling
      iv. Systematic sampling
   b. In order to choose three students to serve as group leaders, all students in a statistics class are assigned a number. Each of those numbers is placed in a hat. Without looking in the hat, the statistics teacher chooses three numbers from the hat.
      i. Simple random sampling
      ii. Stratified random sampling
      iii. Cluster sampling
      iv. Systematic sampling
Appendix E

Student Self-efficacy Beliefs Survey – Post Instruction

1. I believe I will receive an excellent grade on the project related to statistical sampling.
   A. Strongly agree
   B. Agree
   C. Somewhat agree
   D. Neutral
   E. Somewhat disagree
   F. Disagree
   G. Strongly disagree

2. I’m certain I fully understand the most difficult material regarding statistical sampling that was presented in the readings.
   A. Strongly agree
   B. Agree
   C. Somewhat agree
   D. Neutral
   E. Somewhat disagree
   F. Disagree
   G. Strongly disagree

3. I’m confident I understand the basic concepts regarding statistical sampling.
   A. Strongly agree
   B. Agree
   C. Somewhat agree
   D. Neutral
   E. Somewhat disagree
   F. Disagree
   G. Strongly disagree
4. I’m confident I understand the most complex material regarding statistical sampling that was presented by the instructor.

A. Strongly agree
B. Agree
C. Somewhat agree
D. Neutral
E. Somewhat disagree
F. Disagree
G. Strongly disagree

5. I’m confident I did an excellent job on the tasks related to statistical sampling.

A. Strongly agree
B. Agree
C. Somewhat agree
D. Neutral
E. Somewhat disagree
F. Disagree
G. Strongly disagree

6. I did well on this unit about statistical sampling.

A. Strongly agree
B. Agree
C. Somewhat agree
D. Neutral
E. Somewhat disagree
F. Disagree
G. Strongly disagree
7. I’m certain I mastered the skills that were taught in this unit about statistical sampling.

A. Strongly agree
B. Agree
C. Somewhat agree
D. Neutral
E. Somewhat disagree
F. Disagree
G. Strongly disagree
Appendix F

Activity 3.1 - CL

1. Choose a sample of M&M’s from your bag.
   
a. What are the different ways in which you can describe your sample?

b. What can the sample you chose tell you about the population M&M’s in the bag?

c. How would the data about your M&M’s differ if you looked at all of the M&M’s in the bag?

d. Discuss how you chose the sample and why you chose this method.
Prior Knowledge

Before beginning this exercise, the instructor and students in this class had discussed the concepts associated with categorical and quantitative variables, various statistical graphs and outliers, experiments and observational studies, scatterplots and correlation, and linear regression. The instructor had also used the terms sample, population, data, and survey.

Goals

- Students will collaborate with one another and use the materials provided
  - to discuss the concept of sample as they choose a sample of M&Ms
  - to consider the concept of population and how this relates to a sample
- Through collaborative discussion, negotiation, and argumentation, students will develop an understanding of
  - what distinguishes a sample from a population
  - how to make accurate inferences from a sample to a population
  - sampling bias

Materials

- One worksheet per group
- One sandwich size Ziploc bag containing ½ cup of plain M&Ms (approximately 37 candies) per group of students
  a. Use a one pound bag of M&Ms from which to scoop the samples
  b. For sanitation reasons, do not touch the M&Ms
  c. Use additional one pound bags of M&Ms as necessary
d. Have enough Ziploc bags of M&Ms so that each student can have their own bag at the end of the class period

**Launch**

Begin by distributing one worksheet and one small Ziploc bag of M&Ms to each group of students. In order to encourage collaboration, it is imperative to give each group only one set of materials. Tell students that the M&Ms in the Ziploc bags came from a larger, one pound bag of M&Ms. In order to reduce contamination, instruct students to leave the M&Ms in the Ziploc bags while completing the task. Instruct students to read the directions and begin working on the task with their partners. Offer the students no further written or verbal directions at this time.

**Explore**

Allow students ample time to discuss and work through the task with their partners. The amount of time necessary will vary for each different class of students. During the time that students are working through the task with their partners, circulate throughout the classroom taking notes, observing, and listening for any misconceptions or periods when students seem to be off task. Listen for students to use the terms sample, population, and data appropriately and for understanding of these concepts. You may hear students discuss the differences between the M&Ms in their Ziploc bag and the M&Ms in the one pound bag. You might see students separate the candies into groups by color. Be keenly aware of the various ways in which students talk about their sample and the larger population. Intervene only as necessary to encourage students in their collaborations or to bring students back on task. Otherwise, allow students to work through the material on their own and engage in behavior and interact in ways that are
conducive to creating an atmosphere of collaborative learning. While observing, be especially watchful for students to productively argue with their peers regarding ideas about which they disagree, support one another to complete the task by offering alternative explanations and encouragement when necessary, effectively negotiate points of contention, and engage in thoughtful discussion in a statistically literate manner, all indications of collaborative learning. Be aware of students’ progress toward completion and understanding of the concepts of sample, population, the differences between the two, and inferring from a sample to a population so that you can bring the class together for the summary phase when appropriate. If it seems that this task is too complex for students to complete in their small, collaborative groups, consider bringing students together to engage in large group collaboration or whole class discussion.

**Summarize**

Once most of the student groups have come to closure on this task regarding samples and gathering sample data, bring the students back together to discuss their ideas, thoughts, experiences, and lingering questions. As much as possible, allow the students to lead this debriefing session during which they can share their ideas, ask questions of their peers, offer feedback for others, and ask any questions they may have. If possible, do not post or directly ask these questions, but instead listen for students to provide answers to the following questions:

1. What is a sample?
2. What is a population?
3. How do you distinguish a sample from a population?
4. What can you infer about a population from a sample?
5. What does sampling bias mean?

6. Is it okay to have sampling bias?

7. When would sampling bias be absent?

While the students are sharing and discussing, formatively assess the students’ understanding and clarify any areas of confusion. If necessary, make use of the questions listed above to incite discussion of these concepts. You might consider asking students to come up with quick ideas for a study and to identify the population and sample for each. This can help students strengthen their understanding of these concepts and the relationship between the two.
2. Suppose you have a list of all students enrolled in STT 1810 at Appalachian State University this semester. You randomly choose 30 statistics students from this list to survey about their attitudes regarding the legalization of marijuana. You would like to use the results to make generalizations about the attitudes of all undergraduate students in the United States. You email the 30 students and ask them to meet in Walker Hall at 6pm on Thursday night to take a short written survey. Thursday at 5pm, you frantically search the internet for 15 questions regarding attitudes about the legalization of marijuana. You put these 15 questions, along with 5 demographic questions (e.g., What is your age?, What is your race/ethnicity?, etc.) on the survey. 24 of the 30 students you emailed show up on Thursday at 6pm and 19 of those 24 agree to complete the survey.

   a. Discuss this scenario in terms of what you learned from the M&M exercise. Use the following terms to guide your discussion:
      i. Sample
      ii. Population
      iii. Inferences
      iv. Census
      v. Sampling bias

   b. What are some of the problems associated with the process of choosing a sample and gathering data described in this scenario?
Prior Knowledge

Before beginning this exercise, the instructor and students in this class had discussed the concepts associated with categorical and quantitative variables, various statistical graphs and outliers, experiments and observational studies, scatterplots and correlation, and linear regression. The instructor had also used the terms sample, population, data, and survey. Through a whole class discussion following the previous task, students had been exposed to the concepts of sample versus population, inferring from a sample to a population, and sampling bias.

Goals

- Using the scenario provided, students will collaborate with one another and use the worksheet provided to understand the concepts of sample, population, inferring from a sample to a population, census, and simple random sample
- Students will begin to understand about some of the difficulties associated with designing a research study, choosing a sample, and gathering data

Materials

- One worksheet per group

Launch

Distribute one worksheet to each group of students. In order to encourage collaboration among the students in the group, it is imperative to give each group only one worksheet. Ask the groups of students to begin the task by discussing this scenario using the terminology they developed and used while working on the M&Ms task, as well as the language used during the previous Summarize phase. Provide no further written or verbal instructions or information at this time.
Explore

Allow students ample time to discuss and work through the task with their partners, but move about the classroom to ensure that students are staying on task. Listen for students to use the terminology related to the concepts they discussed in the previous task, including sample, population, inference, and sampling bias. Take written notes, intently observe, and identify misconceptions or times when students seem to be confused and unable to move forward with the task. Students may begin to talk about the population of Appalachian students and their thoughts on this issue based upon the sample data. You might hear students question whether or not this is an appropriate sampling method to use in this scenario. Intervene only as necessary to encourage students or to clarify confusing concepts or instructions. Allow students to work through the material in order to get to a place where collaborative learning can occur. While observing, make note of times when students discuss the task using correct academic language, argue in a productive manner regarding points of contention, negotiate with one another to come to agreement with their peers, and provide encouragement and varying perspectives to help each other complete the task, thus experiencing positive social interdependence. Ensure that students are using appropriate statistical language when discussing the concepts of choosing a sample and gathering sample data. Listen for students to exhibit statistical literacy by accurately discussing concepts such as sample, population, statistics, census, and bias. Monitor students’ progress toward understanding of these concepts so that you can bring the class together for the summary phase when appropriate. If at any time it seems that students are inordinately struggling to complete
the task, consider bringing the students together for whole class discussion or large group collaboration.

**Summarize**

Once most groups of students have completed the task, bring the students together for a final period of class discussion around the concepts of sample, population, and gathering sample data. Allow the students to guide this discussion, sharing their ideas and strategies for completing the task, questioning and offering feedback to their peers, and making connections among the various topics included in today’s lesson on collecting and using sample data. Try not to ask students direct questions unless it is necessary to keep the discussion on track. Instead, listen for students to discuss possible answers to the following questions:

1. What is the sample in this scenario?
2. What is the population in this scenario?
3. What distinguishes the sample from the population in this scenario?
4. What can you infer about the population of Appalachian students from this sample?
5. How could a simple random sample of 30 students be chosen?
6. If you wanted to conduct a census, how would you go about that?
7. What are some of the issues related to this scenario (e.g., number of participants, survey items, etc.)?

Use this discussion to formatively assess students’ progress and understanding. If necessary, make use of the questions listed above to incite discussion of these concepts.
Appendix G

Activity 3.2 – CL

1. Suppose you survey 1600 residents of Boone to understand how they feel about potentially charging a fee to ride the AppalCART in-town routes (which are currently free of charge). You find that 57% of the 1600 survey respondents are opposed to this potential fee. Using the calculation for the conservative margin of error, how close is this estimate of 57% to the population percentage of Boone residents?
Prior Knowledge

Prior to this exercise, students had discussed and worked with the concepts of sample versus population, inferring from a sample to a population, sampling bias, census, and simple random sample.

Goals

- Using the scenario provided and working collaboratively with their peers, students will begin to understand the concept of margin of error
- Students will consider the difference between a sample statistic and a population parameter and why these values are typically different from one another

Materials

- One worksheet per group

Launch

Begin with a mini-lecture regarding the concept of margin of error. Briefly discuss how and why to compute a margin of error and how this relates to sampling. Next, distribute one worksheet to each group of students. In order to encourage collaboration, it is important to give each group of students only one worksheet. Encourage students to think about and discuss the concept of margin of error, and its relation to samples and populations, with their group.

Explore

As the students begin to work on this first task, circulate about the classroom observing and listening to students’ interactions. Because this task is rather straightforward and computational, take the time to be especially aware of periods of thoughtful discussion during which students exhibit statistical literacy by accurately
discussing the concept of margin of error, negotiation and productive argumentation regarding ideas upon which students may disagree, and positive social interdependence exhibited by students assisting their peers by explaining the concept of margin of error and encouraging different perspectives, all indications that collaborative learning is occurring. While observing, take detailed notes regarding students’ interactions. Make sure students understand the idea of margin of error by listening for accurate use of terms associated with this concept. Clarify any major misconceptions, but otherwise, allow students to work through the task only with the help of their peers. Be keenly aware of the level of difficulty the majority of students are experiencing. If it seems too great, consider bringing the class together for an alternative method of instruction, such as large group collaboration or whole class discussion.

**Summarize**

Once it appears that most of the student groups have completed this task on margin of error, bring the students together to discuss their experiences, ideas, thoughts, and questions. Allow this discussion to be heavily student led. Do not post or directly ask questions unless absolutely necessary. Instead, listen for students to discuss the following questions:

1. What is the margin of error?
2. How does the margin of error relate the sample to the population?
3. Will there always be a margin of error when working with sample data?

Use this time to formatively assess student knowledge and clear up any misunderstandings. If necessary, make use of the questions listed above to incite discussion of these concepts.
2. In an ABC News Poll conducted between January 21, 2000 and January 26, 2000, a random sample of n=1006 adult Americans was asked, “Compared to buying thing by mail order or in a store, do you think that buying things over the internet poses more of a threat to your personal privacy, less of a threat, or about the same threat level?” The percentages that selected each response are as follows:

- More of a threat: 40%
- Less of a threat: 7%
- About the same threat level: 47%
- No opinion: 6%

a. What is the margin of error for this poll?

b. What is a conservative 95% confidence interval for the population proportion that would have responded ‘More of a threat’? Give your answer for the population percentage as well.

c. Write a sentence or two interpreting the interval that you found in part b above.
Prior Knowledge

Before beginning this exercise, students had discussed and worked with the concepts of sample versus population, inferring from a sample to a population, sampling bias, census, simple random sample, and margin of error.

Goals

- Students will work with their peers to strengthen their understanding of the concept of margin of error
- Students will work collaboratively to understand the concept of confidence intervals, why they are useful, and how confidence intervals relate to the ideas of sample and population

Materials

- This information is on the worksheet that was previously distributed to each group of students

Launch

Begin with a mini-lecture about computing a 95% confidence interval. Then ask students to begin work on the second task of this worksheet.

Explore

As the students begin to work on the second task, move throughout the classroom taking notes and observing and listening to students’ interactions. Specifically note periods of discussion during which students exhibit statistical literacy and the appropriate use of academic language associated with the concepts of margin of error and confidence intervals, the use of negotiation skills when students come to a point of contention, times of productive argumentation regarding ideas upon which students disagree, and the
presence of positive social interdependence or support and encouragement of one another. You may hear students discuss the relationship between the margin of error and the sample statistic, and how these relate to the population. Intervene only when it seems that students are confused or far off task. Otherwise, allow students to work through the task with their peers. If it seems that most of the students are inordinately struggling to successfully complete the task, reconsider the balance of instructional methods used. It may be necessary to bring the class together for another mini-lecture, whole class discussion, or large group collaboration to ensure that students understand the material and successfully complete the task.

Summarize

Once the majority of students have completed the task, bring the students together to debrief regarding the concepts of margin of error and confidence intervals. Allow students to discuss their ideas and thoughts, and to pose any lingering questions. Intervene as little as possible during this discussion. Without directly asking students, ensure that they discuss the following questions:

1. What is the margin of error?
2. How does the margin of error relate the sample to the population?
3. What is a confidence interval?
4. How does the confidence interval relate the sample to the population?
5. Which part of the confidence interval is the sample statistic?
6. Which part of the confidence interval is the margin of error?
During this time, formatively assess student understanding and clarify any misconceptions. If necessary, make use of the questions listed above to incite discussion of these concepts.
3. Suppose that a researcher is designing a survey to estimate the proportion of adults in North Carolina who oppose a proposed law that requires all automobile passengers to wear a seat belt.

   a. What would be the conservative margin of error if the researcher randomly sampled 400 North Carolina adults?

   b. What sample size would be needed to provide a margin of error of no more than 2%?
**Prior Knowledge**

Prior to this task, students had discussed and worked with the concepts of sample versus population, inferring from a sample to a population, sampling bias, census, simple random sample, margin of error, and confidence intervals.

**Goals**

- Working collaboratively, students will further strengthen their understanding of the concept of margin of error
- Students will also begin to understand why it is necessary to determine an appropriate sample size and how the concepts of margin of error and sample size are related to one another

**Materials**

- This information is on the worksheet that was previously distributed to each group of students

**Launch**

Begin with a brief mini-lecture regarding the concept of determining and calculating an appropriate sample size. Once it seems that most students are comfortable with this concept, ask students to work with their peers to complete this final task.

**Explore**

While students work on the final task, walk about the classroom observing students’ interactions and taking notes. Especially note when students are engaging in productive argumentation using the terms associated with margin of error and sample size appropriately, negotiation in order to come to an agreed upon understanding of the concepts of margin of error and sample size, deep, thoughtful discussion regarding these
concepts, and helping, encouraging, and supporting each other in ways that promote positive social interdependence. Listen for accurate understanding of the concepts of margin of error and sample size, exhibited by statistically literate discussions. You might hear students discuss the inverse relationship between margin of error and sample size. Interrupt student discussions only if it seems that students need help moving forward with the task. Otherwise, allow students to work through the task with their peers. If at any time it seems that the majority of students are experiencing great difficulty with the material, consider an alternative method of instruction such as whole class discussion or large group collaboration.

**Summarize**

Once most of the students seem to have completed the task, bring the students together to summarize the topics of margin of error and appropriate sample size. Allow students to lead this discussion, expressing their ideas and thoughts and asking any final questions. During this discussion, do not post or directly pose these questions, but listen for students to provide answers to the following questions:

1. Why should I compute the appropriate sample size before beginning a study?
2. If the sample size increases, how does that affect the margin of error?
3. What will happen if I use a sample that is too small? How will this affect the inferences I make about the population?

Use this time to formatively assess student knowledge and understanding and to clarify any misunderstandings regarding the topics related to margin of error and sample size. If necessary, make use of the questions listed above to incite discussion of these concepts.
Appendix H

Activity 3.3 & 3.4 (1) – CL

Consider the following:

You are interested in the studying habits of undergraduate students at ASU. You obtain a list of the names of all students currently enrolled at the university (see sample below) and want to choose 3000 students to survey (enrollment as of Fall 2010 was 15,137 undergraduate students).

<table>
<thead>
<tr>
<th>Student No.</th>
<th>Last Name</th>
<th>First Name</th>
<th>Class</th>
<th>Major</th>
<th>Age</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Andrews</td>
<td>Jackson</td>
<td>Sophomore</td>
<td>Crim. Justice</td>
<td>19</td>
<td>Male</td>
</tr>
<tr>
<td>2</td>
<td>Azul</td>
<td>Sophia</td>
<td>Senior</td>
<td>Pre-nursing</td>
<td>34</td>
<td>Female</td>
</tr>
<tr>
<td>3</td>
<td>Baker</td>
<td>Amos</td>
<td>Freshman</td>
<td>Undecided</td>
<td>19</td>
<td>Male</td>
</tr>
<tr>
<td>15,135</td>
<td>Young</td>
<td>Sarah</td>
<td>Freshman</td>
<td>Political Sci.</td>
<td>22</td>
<td>Female</td>
</tr>
<tr>
<td>15,136</td>
<td>Zahara</td>
<td>Armin</td>
<td>Senior</td>
<td>Mathematics</td>
<td>24</td>
<td>Male</td>
</tr>
<tr>
<td>15,137</td>
<td>Zost</td>
<td>Corey</td>
<td>Junior</td>
<td>Music</td>
<td>39</td>
<td>Male</td>
</tr>
</tbody>
</table>

- Develop 3 different methods for choosing a sample that accurately represents the population with regard to the topic of interest (i.e., studying habits). Describe, in detail, how you would employ each method and choose your sample.

1.

2.

3.
For each of the 3 sampling methods you developed, discuss why you chose each method. This can include the process you used to develop each sampling method, the reasons behind your choices, etc.

1.

2.

3.
Prior Knowledge

Before beginning this task, students had discussed and worked with the concepts of sample versus population, inferring from a sample to a population, sampling bias, census, simple random sample, margin of error, confidence interval, and determining sample size.

Goals

- Students will work collaboratively to understand the concept of sampling methods and developing single stage sampling plans
- Students will understand the sampling methods associated with various well-known sampling plans, such as a simple random sampling, stratified random sampling, cluster sampling, systematic sampling, and convenience sampling
- Students will use their statistical literacy and statistical reasoning skills to provide justifications for the various sampling plans they developed
- Students will begin to understand why the sampling method chosen makes a difference in the outcome of a study

Materials

- One worksheet per group

Launch

Begin by distributing one worksheet to each group of students. Giving each group of students only one worksheet encourages students to work collaboratively with one another. Instruct the students to begin work with their group without further direction at this time.
Explore

As students begin work on this task, move throughout the classroom observing students’ interactions and taking notes. Listen for students to thoughtfully discuss the topic of developing sampling plans, to engage in productive argumentation when they come to points of contention, to use their skills of negotiation, and to support each other and provide their peers with alternate explanations in ways that promote positive social interdependence. Although students may not use the correct name for each sampling method, you might hear students discuss the process of gathering a simple random sample, stratified random sample, cluster sample, systematic sample, or convenience sample. As students work collaboratively with one another to provide justification for the sampling plans they developed, make sure they have a sound understanding of the various concepts associated with developing sampling plans, such as sample versus population, sampling bias, and inferring from a sample to a population. Specifically note when students exhibit statistical literacy and sound statistical reasoning. Intervene only if it seems that students need assistance or further explanation to successfully complete the task. Otherwise, allow students to be challenged as they work through the task with their peers. Because this can be a complex task for students, they may be more challenged than in the past. If this seems to be the case, consider bringing the students together for large group collaboration or whole class discussion as necessary.

Summarize

One it seems that most of the student groups have completed the task, bring the students together to discuss their experiences, ideas, and questions. Allow the students to lead this discussion and use this time to formatively assess students’ understanding of the
concept of sampling plans. Do not use the names of the various sampling methods, but listen for students to describe any of the following:

1. Simple random sample
2. Stratified sample
3. Cluster sample
4. Systematic sample
5. Convenience sample

Try not to ask these questions directly, but listen for students to discuss answers to the following:

1. What is the difference among the various sampling methods described?
2. Why does it matter which sampling method is used?
3. What issues might arise if the chosen sampling plan is not appropriate for the research study?

During this discussion, formatively assess students’ understanding of the various sampling methods, how they differ, and why it is important to choose an appropriate method of sampling. If necessary, make use of the questions listed above to incite discussion of these concepts.
Appendix I

Activity 3.3 & 3.4 (2) – CL

Suppose you want to better understand the voting preferences of registered voters in the state of North Carolina. Develop ONE sampling plan that includes at least TWO of the following probability sampling methods. That is, develop a multistage sampling plan to survey the registered voters in North Carolina. An NC county map is provided below should you need it. Justify why you chose your specific sampling plan.

1. Stratified Random Sampling:
   a. Divide the population into strata, or subgroups that represent the population with regard to the variables to be measured.
   b. Choose a simple random sample from each strata and this becomes your overall sample.
   c. Survey the chosen units (e.g., people, animals, M&Ms, etc.).

2. Cluster Sampling:
   a. Divide the population into clusters, or subgroups that represent the population with regard to the variables to be measured.
   b. Choose a simple random sample of clusters and every member of the chosen clusters is included in your sample.
   c. Survey the chosen units (e.g., people, animals, M&Ms, etc.).

3. Systematic Sampling:
   a. Randomly choose a starting point on the full list of population units (also called the sampling frame) and then choose every k\textsuperscript{th} unit on the list.
   b. Survey the chosen units (e.g., people, animals, M&Ms, etc.).
Prior Knowledge

Before beginning this task, students had discussed and worked with the concepts of sample versus population, inferring from a sample to a population, sampling bias, census, simple random sample, margin of error, confidence interval, determining sample size, single stage sampling plans such as stratified random sampling, cluster sampling, systematic sampling, and convenience sampling.

Goals

- Students will work collaboratively to understand the concept of developing multistage sampling plans
- Students will use their statistical literacy and statistical reasoning skills to provide justification for the multistage sampling plan they developed

Materials

- One worksheet per group

Launch

Begin by distributing one worksheet to each group of students. The purpose of giving each group of students only one worksheet is to encourage collaboration among the students. Remind students of the various sampling plans they developed and discussed during the previous class period. Ask the groups of students to begin working on this task without further instruction at this time.

Explore

As students begin to work on this task, circulate through the classroom observing, listening, and noting students’ interactions. Listen for thoughtful discussion using the language of developing sampling plans, productive argumentation among students as
they develop a multistage sampling plan, effective negotiation of ideas, and indications of positive social interdependence such as encouragement, providing alternative perspectives for certain concepts, and general support of one another. Ensure that students understand the concept of multistage sampling by listening for appropriate discussion of combining the sampling methods of simple random sampling, stratified random sampling, cluster sampling, and systematic sampling. Interrupt student groups only if it seems that they would not be able to proceed otherwise. Keep in mind that this task is rather complex for non-expert statistics students, so allow them to be challenged as they work through the task with their peers. If it seems that several groups find it difficult to successfully navigate the task, consider bringing the students together for a mini-lecture, large group collaboration, or whole class discussion. It is imperative, however, that students are provided ample time to experience challenge by the task before coming together as a whole class.

**Summarize**

Once most of the students have completed the majority of the task, bring them together to discuss their strategies for creating a multistage sampling plan, their difficulties, and their remaining questions. If possible, allow this discussion to be led by the students. Without posting questions or asking students directly, listen for students to discuss the following:

1. When is it appropriate to use a single stage sampling plan and when it is appropriate to use a multistage sampling plan?
2. For a given scenario, what is the most appropriate combination of single stage sampling plans that will allow me to choose a sample that best represents the population?

3. What are some issues that might arise if I choose an inappropriate sampling plan? During this time, formatively assess the understanding and knowledge of the students and provide any clarifications necessary. If necessary, make use of the questions listed above to incite discussion of these concepts.
Appendix J

Activity 3.5 & 3.6 – CL

1. Together with your partner, choose a topic you would like to know more about.

2. Write 1 overall question you would like to answer about this topic. Make sure your question includes the population to which you would like to generalize your sample results.

   For example, let’s say I’d like to know more about people’s beliefs about religion. My overall question might be, “How do the religious beliefs of adults in the U.S. differ by region (i.e., the South, the Midwest, the Northeast, the Northwest, etc.)”? In this scenario, ‘adults in the U.S.’ is the population to which I’d like to generalize my sample results.

3. Write 5 survey questions that you will administer to a sample of participants. Make sure these 5 questions will provide you with the data you need to answer your overall question above. Think about bias and other potential issues when crafting these questions.

1.

2.

3.

4.

5.
Prior Knowledge

Before beginning this task, students had discussed and worked with the concepts of sample versus population, inferring from a sample to a population, sampling bias, census, simple random sample, margin of error, confidence interval, determining sample size, single stage sampling plans such as stratified random sampling, cluster sampling, systematic sampling, convenience sampling, and multistage sampling plans.

Goals

- Students will begin to recognize some of the issues that can arise when developing a research study, such as selection bias, response bias, and nonresponse bias
- Students will become aware of the difficulty of writing appropriately worded survey items and the issues that can arise if survey items are worded such that they do not gather data to answer the questions of interest

Materials

- One worksheet per group

Launch

Begin by distributing one worksheet to each group of students. By giving each group of students only one worksheet, students are more likely to engage in collaborative learning. Ask the groups of students to begin working on this task without further instruction at this time.

Explore

As students begin to work on this task, move about the classroom listening to and observing students’ interactions. Listen for students’ use of the statistical language they
have learned over the course of this instructional unit, including sample, population, statistics, inference, bias, simple random sample, stratified random sample, cluster sample, and systematic sample. Observe for periods of creative, thoughtful discussion including these statistical terms, negotiation of various points of contention and coming to an agreed upon understanding, effective argumentation of points upon which students disagree, and students supporting and helping one another complete the task by offering alternative explanations of confusing concepts. Make note of the strategies students use to develop their topic, overall question, and five associated survey questions, as these strategies are often demonstrations of students’ statistical literacy and statistical reasoning skills. You might hear students express concern with the way in which a particular survey item is worded. Make sure students understand the concepts associated with developing a research idea and creating survey items, including many of the ideas discussed throughout this instructional unit. Intervene only if it seems that students are at an impasse and would not be able to proceed otherwise. Because this task requires students to think on their own to develop their topic and related questions, it can be more challenging than the other tasks they completed during this instructional unit. If you see that several groups find this task inordinately challenging, consider bringing the students together for large group collaboration or whole class discussion. In this case, a mini-lecture is probably not an appropriate alternative. Keep in mind that struggling with a concept or task can be beneficial to students and that this time of cognitive disequilibrium can help students organize their thoughts and develop ideas they may not otherwise.
Summarize

Once most of the students have completed the majority of the task, bring them together to discuss their strategies for deciding upon a topic, determining an overall question, and developing survey questions. Allow students to express their difficulties with this task as well as any lingering ideas or questions. As much as possible, allow this to be large group collaborative session during which students pose questions to their peers and offer one another constructive feedback. Listen for students to discuss answers to the following questions:

1. Why does the wording of a survey item matter?
2. How do you ensure that sample survey participants best represent the population?
3. How can you word survey items to avoid making people feel uncomfortable?
4. What can you do if people do not want to participate in a survey or experiment?
5. What can you do if, once you’ve started administering a survey, you realize that the survey items are not actually answering the questions of interest?

Use this time to formatively assess students’ understanding and knowledge and answer any final questions students may have. If necessary, you can bring the concepts of this instructional unit regarding statistical sampling together by asking students to discuss their topic, overall question, and five survey items considering the concepts of sample, population, data collection, making inferences from a sample to a population, margin of error, various forms of bias, confidence intervals, and single stage and multistage sampling plans. If necessary, make use of the questions listed above to incite discussion of some of these concepts.
REFERENCES


