12-2017

The Influence of Previous Subject Experience on Interactions During Peer Instruction in an Introductory Physics Course: A Mixed Methods Analysis

Judy A. Vondruska
University of Nebraska-Lincoln, judy.vondruska@sdstate.edu

THE INFLUENCE OF PREVIOUS SUBJECT EXPERIENCE ON INTERACTIONS DURING PEER INSTRUCTION IN AN INTRODUCTORY PHYSICS COURSE: A MIXED METHODS ANALYSIS

by

Judy A. Vondruska

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy

Major: Educational Studies

Under the Supervision of Professor Allen Steckelberg

Lincoln, Nebraska

December, 2017

THE INFLUENCE OF PREVIOUS SUBJECT EXPERIENCE ON INTERACTIONS DURING PEER INSTRUCTION IN AN INTRODUCTORY PHYSICS COURSE: A MIXED METHODS ANALYSIS

Judy A. Vondruska, Ph.D.

University of Nebraska, 2017

Advisor: Allen Steckelberg

Over the past decade, peer instruction and the introduction of student response systems has provided a means of improving student engagement and achievement in large-lecture settings. While the nature of the student discourse occurring during peer instruction is less understood, existing studies have shown student ideas about the subject, extraneous cues, and confidence level appear to matter in the student-student discourse. Using a mixed methods research design, this study examined the influence of previous subject experience on peer instruction in an introductory, one-semester Survey of Physics course. Quantitative results indicated students in discussion pairs where both had previous subject experience were more likely to answer clicker question correctly both before and after peer discussion compared to student groups where neither partner had previous subject experience. Students in mixed discussion pairs were not statistically different in correct response rates from the other pairings. There was no statistically significant difference between the experience pairs on unit exam scores or the Peer Instruction Partner Survey. Although there was a statistically significant difference between the pre-MPEX and post-MPEX scores, there was no difference between the members of the various subject experience peer discussion pairs. The qualitative study, conducted after the quantitative study, helped to inform the quantitative results by

exploring the nature of the peer interactions through survey questions and a series of focus groups discussions. While the majority of participants described a benefit to the use of clickers in the lecture, their experience with their discussion partners varied. Students with previous subject experience tended to describe peer instruction more positively than students who did not have previous subject experience, regardless of the experience level of their partner. They were also more likely to report favorable levels of comfort with the peer instruction experience. Students with no previous subject experience were more likely to describe a level of discomfort being assigned a stranger for a discussion partner and were more likely to report communication issues with their partner. Most group members, regardless of previous subject experience, related deeper discussions occurring when partners did not initially have the same answer to the clicker questions.

ACKNOWLEDGEMENTS

The formidable task of completing a dissertation does not occur without sacrifice and support. I would like to acknowledge and thank each member of my committee for their insight and support in this endeavor. Through this journey many changes occurred with the greatest being the untimely death of my initial advisor, Dr. Delwyn Harnisch. Del was a man of strong belief who lived his life accordingly. I admired his convictions and appreciated his kind advice.

I want to thank Dr. Allen Steckelberg, the only remaining member of my original committee, for stepping up to become my advisor after Del's passing. I have enjoyed working with you and greatly appreciated your insights into improving my work. I want to also thank Drs. Lawrence Scharmann, Julie Thomas and Douglas Golick for agreeing to serve on my dissertation committee after previous members were unavailable due to sabbatical or retirement. I realize the incredible time commitment each of you already had and truly appreciate your willingness to take on one more doctoral student. I appreciate your advice and expertise in helping me create a study worth conducting.

I want to thank my husband, Dr. Stephen Krebsbach, for all his support over these years completing this doctoral program. I want to thank him most of all for helping me get out of my own way to finally get this done. There have been many sacrifices in pursuit of this dream and I look forward to many years of spending time with you golfing, traveling, and watching baseball and football.

I want to thank the Physics Department at South Dakota State University for allowing me professional leave when needed to meet with my advisor or committee.

I want to thank all my friends and family who have supported me over the years with their encouraging words. I want to thank my family members who are no longer here to share in this dream but who supported me through it - my Aunt Dorothy and Uncle Tom who were both inspirational educators and especially my Mom, Beverly B. Vondruska (Olson), who always believed in and supported me.

I want to dedicate this dissertation to my Dad, Joseph E. Vondruska, who inspired in me the love of science and teaching. Dad was the first in his family to attend college and achieved a master's degree in chemistry. He taught high school science for over 40 years, most of it in Schuyler, Nebraska, inspiring generations of students. I am very grateful he is still here to share in this accomplishment with me.

With gratitude,

Judy

TABLE OF CONTENTS

Chapter 1: Introduction 1

Research Problem 1

Purpose of Study 4

Research Questions 6

Phase I – Quantitative Questions 6

Phase II – Qualitative Questions 7

Phase III – Mixed Methods Questions 8

Definition of Terms 8

Significance of Study 9

Limitations 9

Chapter 2: Literature Review 11

Increasing Student Engagement 11

Attitudinal and Motivational Impacts 13

Impact on Student Learning 16

Varying Levels of Effectiveness 18

Methods of Implementation 20

Feedback 21

Nature of Discussions 22

Summary 25

Chapter 3: Research Design and Methodology 27

Mixed Methods Research Design 27

Philosophical Assumptions 28

Explanatory Sequential Design 28

Target Population and Sampling 30

Data Collection and Analysis 30

Validity and Reliability 43

Instrumentation 45

Ethical Considerations 51

Summary 51

Chapter 4 – Quantitative and Qualitative Results 52

Phase I – Quantitative Results 52

Phase II – Qualitative Results 66

Chapter 5 – Discussion 81

Development of Qualitative Questions from Quantitative Results 82

Using Qualitative Results to Explain Quantitative Results 84

Implications 87

Limitations 89

Recommendations for Future Research 90

Appendix A – IRB Approval Letters 92

Appendix B – Qualitative Interview Documents 94

Appendix C – Peer Instruction Partner Survey 97

Appendix D – Word Clouds 98

References 100

CHAPTER 1: INTRODUCTION

Large lecture courses are common in universities especially at the introductory or general education level. Large-lecture classrooms are physically designed with a “stand and deliver” orientation positioning the instructor at the front of the room and students in fixed seating arrangements - as an audience. Although increased student engagement through student-student and student-teacher interactions has been shown to improve learning in the classroom (Chickering and Gamson, 1987; Kim, K., Sharma, P., 2013; Prather, E., Rudolph, A., et.al., 2009), the large-lecture class size and classroom orientation diminishes the ability of the instructor to engage with students in a meaningful way. Over the past decade, peer instruction and the introduction of student response systems (SRS), sometimes referred to as "clickers", has provided a means of improving student engagement in large-lecture settings.

Through the use of clickers, instructors can pause periodically during a lecture to ask questions about the subject matter being covered. Students provide their responses with the SRS after which the aggregate results are displayed for the class. The use of the SRS not only provides engagement and interaction but also feedback to both students and the instructor and feedback has been shown to have a significant impact on student achievement (Hattie, 1982). The value of feedback using SRS was also evidenced in a study by Doucet, Vrins, and Harvey (2009) where nearly 100% of students "indicated that the ability to evaluate their own strengths and weaknesses was the single most useful advantage of using the ARS [sic... audience response system]." The feedback provided by the aggregate SRS data can be used by an instructor to provide remediation during class, ask additional questions to further clarify misunderstandings, or comfortably move forward with new material.

Initial research into the use of SRS focused primarily on engagement and attitudes. Survey results across a variety of classroom sizes and subject matter demonstrated that students found the use of clickers beneficial in their learning (Beekes, 2006; Yourstone, Kraye & Albaum, 2008; Salemi, 2009; Stowell & Nelson, 2007; Trees & Jackson 2007). Factors contributing to positive student perception included a desire to be involved, devaluing of lecture, value of class feedback, class standing, previous experience with lecture courses, anticipated course performance, and amount of clicker use in the classroom (Trees & Jackson, 2007).

More recent research into SRS use has focused on its impact on student learning. In a study of student attitudes and conceptual understanding in a general chemistry course, the use of SRS was not only viewed as helpful by students but that belief was supported by exam results (Donovan, 2008). Shapiro and Gordon (2012) found the use of clicker questions improved student performance on exam questions as well or better than explicitly telling students the information would be on the test. Blood (2012) found performance on quiz questions used to test for intermediate and long-term recall were significantly improved with the use of SRS. These types of results have been supported by numerous other studies on the impact of SRS on student learning (Mayer, et. al., 2009, Pradhan, 2005, Kennedy & Cutts, 2005).

Whether the improvement in learning is related to the use of SRS or the just the pedagogical practice of stopping and asking thought-provoking questions during lecture is debatable. Some researchers suggest the nature of the questions may be the underlying factor in clicker success (Christopherson, 2011; Price, DeLeone, & Lasry, 2010; Gray & Steer, 2012). Questions designed to direct students' attention, stimulate specific cognitive processes, communicate information to the student and instructor, and facilitate student ability to articulate and confront key ideas within the content are most effective (Beatty, Gerace, Leonard, & Dufresne, 2006).

As part of the SRS process, instructors may have students engage in a discussion with one or more classmates to determine a correct answer to the question. This process is referred to as peer instruction (Mazur, 1997). The exact methodology used during peer instruction can vary including the timing and nature of the discussion. Regardless of the specific process used, discussion with classmates has been shown to increase the percent of students ultimately selecting the correct answer on the question, especially when the initial percentage of correct answers is around 50% (Crouch and Mazur, 2001). Peer instruction, which includes questions focusing on student misconceptions, has been found to improve conceptual and quantitative understanding of physics material in both introductory calculus and algebra-based physics courses (Crouch and Mazur, 2001).

Improvement in student learning can also be related to the feedback provided by the instructor as a result of student responses (Yourstone, et.al., 2008). In a study by Smith, Wood, Krauter and Knight (2011) clicker questions were presented with three formats - peer discussion only, listening to instructor explanations only, or with peer discussion followed by instructor explanation. Results showed peer instruction followed by instructor explanation did more to improve overall student performance than either of the other two formats.

Recent SRS research has begun to focus on the nature of the discourse occurring during the peer instruction interaction. Recorded discussions during peer instruction have demonstrated three broad categories of "non-standard" (not focused on the specific question detractors or question concepts) conversations: 1) student ideas about the subject that were unanticipated by the question answers; 2) statistical feedback misrepresenting student understanding such as extraneous cues or the confidence of the partner; and, 3) conversation pitfalls related to students ability to engage in conversation effectively (James and Willoughby, 2011). These "non-standard" conversations may have an impact on the success of peer instruction for some students.

Research Problem

If the conversation positively impacts the nature of the response, then optimizing that interaction is of value. Since student ideas about the subject, extraneous cues and confidence level appear to matter in the student-student discourse, a student's previous experience with the subject matter may be a factor in the success of peer instruction. Previous subject experience has already been shown to impact academic success in the college physics classroom (Sadler and Tai, 2001) and with the number of students enrolling in high school physics having more than doubled since 1987 (American Institute of Physics, 2014), it is fair to assume a significant number of students in an introductory college physics course will have had high school physics, although not necessarily a majority. This varied background in previous subject matter experience will result in discussion partners during peer instruction with different levels of subject matter expertise. The previous subject experience of a student may influence attitudes about physics, the nature of the student-student interactions during peer instruction and the correct response rates measured with SRS, providing the instructor with a false measure of the level of learning occurring in the classroom. Exploring the influence of previous subject experience on interactions during peer instruction is needed to better understand its level of success for all students in introductory college physics.

Purpose of Study

The purpose of this study was to explore the influence of previous subject matter experience on learning, attitudes and interactions of students using peer instruction in a one-semester, college-level, Survey of Physics course. An explanatory, sequential, mixed methods design was used which involved collecting quantitative data first and then explaining the quantitative results with in-depth qualitative data.

In the first, quantitative phase of the study, students were assigned to one of three subject experience pair (SEP) discussion groups based on whether or not they had high school physics. Students who had high school physics were designated ‘P’ and students who did not have high school physics were designated ‘N’. Students were assigned one discussion partner for peer instruction, resulting in three types of subject experience pairs for peer instruction - PP, PN, and NN. Responses to clicker questions asked during lecture utilizing the peer instruction technique were collected and compared. This included responses prior to and after partner discussion. Additional quantitative data collected and analyzed included student exam scores, student ratings on the Maryland Physics Expectations Survey (MPEX), and results on a peer instruction partner survey (PIPS).

In the second, qualitative phase, a subset of participants from each SEP was selected to participate in one of four focus groups (P from PP, P from PN, N from PN and

N from NN) where details about participant experiences and perceptions about peer instruction were evaluated. Data from the qualitative phase was used to provide a better understanding of the quantitative data as well as provide a more in-depth understanding of the nature of the interactions between the various SEP.

Research Questions

As a mixed methods design there are three categories of research questions – quantitative, qualitative and mixed.

Phase I - Quantitative Questions

1. What is the influence of previous subject experience on SRS correct response rates prior to peer discussion?
2. What is the influence between previous subject experience and SRS correct response rates after peer discussion for each SEP?
3. What is the influence of previous subject experience and peer instruction on student understanding of the subject matter as measured on unit exams?
4. What is the influence of previous subject experience and peer instruction on student perceptions about their discussion partner, the nature of physics and how to be successful in a physics course?

   The specific sub-questions for Phase I are:

   1. How do correct first response rates compare between the P and N groups?
   2. How do correct response rates of each SEP compare after partner discussion?
   3. How do students in each SEP perceive the nature of physics and how to learn physics, as measured on the MPEX Survey?

4. How does the achievement on unit exam scores compare between each SEP?
5. How do attitudes about partners compare between each SEP?

Phase II - Qualitative Questions

For the second, qualitative phase of this study research questions focused on the nature of previous experience with clickers and peer instruction, participant perceived value in the process of peer instruction and the nature of the peer interactions during discussion of clicker questions:

1. What is the nature of previous clicker use or peer instruction experience in each SEP? How might this previous experience have influenced the peer instruction and the quantitative results?
2. How do students in each SEP perceive the value of clickers in learning physics and how might their perceptions explain influence the nature of peer instruction and the quantitative results?
3. How do students in each SEP perceive the value of peer instruction in learning physics and how might their perceptions explain the quantitative results?
4. How do students in each SEP describe the nature of discussion during peer instruction and how might their perceptions explain the quantitative results?

The research sub-questions for Phase II were refined and expanded based on the results of the first, quantitative phase of the study.

Phase III - Mixed Methods Question

What is the difference in experiences in a college Survey of Physics course between peer instruction discussion partners who have no previous subject experience compared to student discussion pairs in which only one or both students have previous subject experience?

Definitions of Terms

For the purpose of this study, the following definitions apply:

Previous Subject Experience is a designation to describe whether a student had high school physics. A designation of 'P' will be used to describe a student who had high school physics and a designation of 'N' will be used to describe a student who did not have high school physics.

Subject Experience Pairs (SEP) refer to the level of previous subject experience of discussion partners assigned for this study. PP refers to a student pair in which both students have taken high school physics. PN is a student pair in which one student had high school physics and one did not. NN refers to a student pair where neither student had high school physics.

Peer Instruction is an instructional methodology in which lectures are interspersed with conceptual questions, called *ConcepTests*, designed to expose common difficulties in understanding the material (Mazur, 1997).

Student Response Systems (SRS) refer to any electronic device with which students can provide individual answers to instructor questions. This type of system is also referenced in the literature as an audience response system (ARS), personal response systems (PRS), or clicker.

Maryland Physics Expectation Survey (*MPEX)* was developed at the University of Washington in the fall of 1992 (Redish, Steinburg, Saul, 1998) as a means of assessing students' cognitive expectations of a physics course as related to their understanding of the process of learning physics and the structure of physics knowledge.

Statistically Significant is a result that is very unlikely to occur when the null hypothesis is true. The result is sufficient to reject the null hypothesis (Gravetter & Wallnau, 2007).

Significance of Study

This research study informs the practice of peer instruction in an introductory physics course. The findings from this research illustrate the influence of previous subject experience on correct response rates during peer instruction, students' perceptions about peer instruction and physics, and student success in learning physics. The findings from this research study are of interest to educators utilizing peer instruction and researchers investigating the effectiveness of peer instruction.

Limitations

The study is limited in its findings in the following ways:

1. During the quantitative phase, there was the potential for non-response in all forms of data but student absences during clicker data collection is most significant. There was a graded attendance component in the course which should kept this non-response rate small. Students absent during times in which the MPEX was administered were asked to complete it at another time, close in date to the original time taken by the rest of the class.

2. During the qualitative phase, only a sampled subset of each SEP was invited to participate in the focus groups. These students may not be entirely representative of the subject experience group as a whole.
3. There is always a level of interpretation in coding and development of themes in qualitative research. Alternative interpretations of the qualitative data may be made by different researchers.
4. A stratified sampling technique is used to assign students to an SEP. This was an effort to equally distribute students into SEP based on previous subject experience, gender and year in college. Even with this stratified sampling technique, this study occured within a single one-semester Survey of Physics course for non-science majors, a sample of convenience. This may limit the generalizability of the study.

CHAPTER 2: LITERATURE REVIEW

Teaching large college lecture sections can be a daunting task. Beyond just the numbers associated with grading, there is a more profound pedagogical teaching question – how to facilitate interactions and effective learning. Active learning methods which increase student-student and student-instructor interaction have been established as an effective method for learning in the classroom (Chickering and Gamson, 1987; Kim, K., Sharma, P., 2013; Prather, E., Rudolph, A., et.al., 2009). In a group study of 4000 students enrolled in 69 sections of courses taught by 36 different instructors at 31 institutions gains in learning on the topic of light and spectroscopy in astronomy courses was due to the nature of the instruction in the classroom (active-engagement, learner-centered) not on the type of class or the institution (Prather, 2009). The study also found higher interactivity resulted in the largest learning gains. Increased interactions inside and outside of class, especially the type that includes some type of feedback have been developed and integrated into classrooms over the past several decades and have demonstrated increases in student learning. These techniques have included tutorials (McDermott, L., Shaffer, P., et al., 2002), case-based learning (Ciraj, A.M., Vinod, P., and Ramnarayan, K., 2010), peer instruction (Mazur, E., 1997), socratic homework systems (Morote, E.S., Pritchard, D.E., 2009), multiple representations of concepts and principles (Brewe, E., 2008), and restructuring of the physical learning environment (Oliver-Hoyo, M. and Beichner, R., 2004).

Increasing Student Engagement

While group work and student-teacher interaction in small classes is relatively easy to facilitate, it is more difficult in larger classes due not only to the number of students but also the physical structure of the classroom. An instructor cannot possibly interact individually with hundreds of students in a given lecture time period nor can the instructor physically position themselves near most students to hear or participate in discussions.

One method developed to increase student participation and student-teacher interaction in the classroom is the use of some type of student response system (Mazur, E., 1997; Wieman & Perkins, 2005; Stowell & Nelson, 2007; Blasco-Arcas, et.al., 2013). Initial introduction of these methods in physics classrooms included the use of a coded (A,B, C, D, etc.) system on flashcards (Mazur, 1997) but now also includes the current wide-spread use of electronic student response systems (SRS) often referred to as clickers. Clickers systems have now evolved to include the ability of students to enter responses using personal cell phones, laptops, and tablets.

The use of SRS allows instructors to ask questions of large groups of students and acquire aggregate responses which can be shared live with the class. Individual responses can also be used, if the clickers have been registered with the instructor, to evaluate individual responses at a later time. During class, clicker responses provide feedback to both instructors and students, allowing students to gain immediate feedback about their level of understanding of the material and the instructor to use the responses to gauge the speed of the lecture and determine the need for re-teaching or modifying the lecture on the spot (Kolikant, Drane, & Calkins, 2010; Sevian & Robinson, 2011). Instructors can also use the results from clicker questions to identify the “muddiest points” at the end of a lecture (King, 2011) and to modify future lectures, dropping material that is understood by the majority of the class (Anderson, Healy, Kole, & Bourne, 2011).

Attitudinal and Motivational Impacts

During the first several years of SRS use, research primarily focused on the attitudinal and motivational impact on students or more generally, the concept of student engagement. Survey results in a variety of classrooms demonstrated that students found the use of clickers beneficial in their learning (Bazen, & Clark, 2005; Beekes, 2006; Blood & Neel, 2008; Yourstone, Kraye & Albaum, 2008; Salemi, 2009; Stowell & Nelson, 2007; Trees & Jackson 2007;). The impact of SRS appears to cross content disciplines as well as grade level. A research study at the University of Carolina looked at the student and faculty impact of SRS implementation in science, math, engineering, art, education, and foreign languages (Crews, Ducate, Rathel, Heid, & Bishoff, 2011). Faculty in the study cited four primary uses of the SRS - encourage classroom participation, facilitate group discussion, assess students' mastery of content, and to help students learn new terms. Student survey results indicated perceived high clicker effectiveness for attendance, student engagement, instructor feedback, increase in class discussion, and increase in student learning. The student survey results were in agreement with faculty observations and course assessments. Similar results have been found in studies in courses on financial accounting (Premuroso, Tong, Beed, 2011; Carnaghan, et.al., 2011; Beekes, 2006; Carnaghan & Webb, 2007), psychology (Shaffer and Collura, 2009; Stowell & Nelson, 2007; Morling, McAuliffe, Cohen, DiLorenzo, 2008; Dallaire, 2011), physics (Mazur, 1997; Corpuz, Corpuz, Rosalez, 2010; Roxas, Carrion-Monterola & Monterola, 2010; Keller, et. al., 2007), astronomy (Prather, E. E., Rudolph, A. L., Brissenden, G., & Schlingman, 2009), mathematics (d'Invero, Davis, White, 2003; Ray, Hugh, & Su, 2003), social work (Quinn, 2010), upper-division physics (Perkins & Turpen, 2009), health sciences (FitzPatrick, Finn, Campisi, 2011), nursing (Hunter, Revell & McCurry, 2010; Mordhorst, 2010; Porter & Tousman, 2010; Skiba, 2006; Sternberger, 2012), biology (Hunter, Rulfs, Caron, Buckholt, 2010; Wolter, Lundeberg, Kang, Herreid, 2011; Blasco-Arcas, et.al., 2013; Lundeberg, Kang, Wolter, DelMas, Armstrong, Borsari, Hagley, 2011), business (Eastman, Iyer, Eastman, 2011), economics (Salemi, 2009; Ghosh & Renna, 2009), family and consumer sciences (Gentry, 2007), human development (Beckert, Fauth, & Olsen, 2009), forestry (Bibles, 2011), management (Keough, 2012), marketing (Lincoln, 2008), and political science (Gormley-heenan & McCartan, 2009).

While some of this positive attitude may be attributable to simply the uniqueness of the SRS system (a “newness” or novelty effect), studies have shown student attitudes can be related to how the SRS is used in the classroom. Trees & Jackson (2007) surveyed 1500 students in three university departments and found several factors which contributed to positive student perception of SRS. These included a desire to be involved, devaluing of lecture, value of class feedback, class standing, previous experience with lecture courses, anticipated course performance, and amount of clicker use in the classroom.

Positive attitudes and motivation for the use of SRS may be related to student attention. In a study of three levels of general chemistry, students self-reported less attention decline during lecture with the use of SRS than without (Bunce, Flens, & Neiles, 2010). In a study of first-year pharmacy students, 98% reported that SRS questions, placed strategically during lecture, helped them maintain attention (Cain, Black, & Rohr, 2009).

The use of student response systems can facilitate discussions of personal and controversial topics as well. Stowell, Oldham and Bennett (2010) found the use of electronic SRS reduced conformity and shyness in a psychology course where numerous controversial questions were asked of students. Students in this study were also more likely to respond to a given question with an SRS than if they had been instructed to raise their hands. Micheletto (2011) found similar results in a Principles of Marketing course. The use of SRS to engage students in discussions about sensitive and controversial topics related to business led to significant student reflection on the topic and changes in perceptions and attitudes. SRS have also been used in psychology courses to replicate known empirical phenomenon which would otherwise have been affected by student open responses, discussions, or hand-raising (Cleary, 2008).

Positive perception of SRS use is not universal, however. Kang, et. al. (2011) found more positive attitudes toward clicker use from women and non-science majors in introductory biology courses. Students in undergraduate science courses found clickers more useful in classes in which there was active participation through discussion (Keller, Finkelstein, et.al., 2007). In situations where group response systems have been used rather than individual response systems, declines in engagement have been found as well as little or no difference in student satisfaction or learning (Carnaghan & Webb, 2007). A student's year in school appears to play a role in the value given active learning techniques in the classroom. Welsh (2012) found fourth and fifth-year undergraduate students more likely to view in-class active learning techniques as a waste of time compared to third-year students. Women were also more likely to find active learning techniques as significantly important to improving their understanding of the material and providing interactions with other students and the instructor.

Impact on Student Learning

Although student surveys about the value of SRS in learning content often report high student perception about the value of SRS, studies to actually measure this increased learning effect have been less numerous than attitudinal studies. For those studies focusing on the impact of SRS on learning, perception does appear to be reality. In a study of students in a college-level psychology course, students who used clickers scored statistically significantly higher on course exams (1/3 of a grade point) when they had used clickers with in-class questions compared to an identical class in which the same questions had been presented with no clicker use and another class in which in-class questions were never provided (Mayer, et. al., 2009). In a study of student attitudes and improvement in conceptual understanding in a general chemistry course, the use of SRS was not only viewed as helpful by students but that belief was supported by exam results (Donovan, 2008). These improvements in learning with SRS are not just due to attention-grabbing. Shapiro and Gordon (2012) found the use of clicker questions improved student performance on exam questions as well or better than explicitly telling students the information would be on the test. Blood (2012) found performance on quiz questions used to test for intermediate and long-term recall were significantly improved with the use of SRS although no association of SRS and improved student engagement was found.

In a randomized block experimental design of four sections of Operations Management, Yourstone, et.al. (2008) found the use of clickers during class, in conjunction with immediate instructor feedback on the questions, had a positive and statistically significant impact on student learning as measured on test scores as compared to students who did not use clickers and received feedback on questions at a later time. In a randomized, controlled trial of 17 obstetrics and gynocology residents, Pradhan (2005) found a 21% improvement between pretest and posttest scores for residents who used clickers during lecture as compared to a 2% improvement for residents who did not use clickers. Kennedy & Cutts (2005) found a positive association between the use of electronic voting systems and achievement of students on learning outcomes in a first-year computer science course. In a study of the use of SRS in managerial courses, students in the SRS courses performed, on average, 3.5 percentage points better than students in the non-SRS courses. The study also found the use of SRS helped low-GPA students more than high-GPA students (Edmonds & Edmonds, 2008).

Crouch and Mazur (2001) reported data from ten years of experience with peer instruction (including SRS) showing students in calculus- and algebra-based introductory physics courses for non-majors performed better on conceptual questions and quantitative problems than those students who did not experience peer instruction. Fagen, Crouch, and Mazur (2002) collected data from 384 teachers using peer instruction (and SRS) in high schools, community colleges, 2- and 4-year colleges, and found more than 80% reported successful implementation and positive learning gains using peer instruction.

Some researchers suggest the nature of the questions may be the underlying factor in clicker success (Christopherson, 2011; Price, DeLeone, & Lasry, 2010; Gray & Steer, 2012). Effective questions have an explicit pedagogic purpose and include content, process and metacognitive goals. Questions designed to direct students' attention, stimulate specific cognitive processes, communicate information to the student and instructor, and facilitate student ability to articulate and confront key ideas within the content are most effective (Beatty, Gerace, Leonard, & Dufresne, 2005). The importance of the question is also supported by Anthis (2011) who isolated the effect of clickers and demonstrated more positive effects on exam scores from students raising their hand to answer pedagogic questions compared to the students answering the same questions with SRS. Grouping questions can also be beneficial. Rather than asking one question per concept, a sequence of SRS questions on the same concept has been shown to be more effective in helping students learn physics (Lee, Lin, Reay, & Lei, 2011). Clearly the pedagogical practice has significance in the success of SRS in the classroom.

Varying Levels of Effectiveness

The impact of SRS use on learning is not always positive nor does it always have a lasting effect. Karaman (2011) found that while students who used audience response systems in class had significantly higher scores on assessments compared to students who verbally responded to questions in class, the effect was only measureable in the first four weeks in the course. After this time, there was no difference measured in the two groups. In a study of a variety of health sciences courses at various levels, Fitzpatrick, Finn, and Campisi (2011) found that although students found clickers useful to a large degree in all courses assessed, the use of clickers did not always make a difference in student quiz/exam scores. Research results comparing two years of non-clicker use with two years of clicker use in the same classes showed that clicker use improved scores in lower-level courses but not upper-division courses. A study of the effectiveness of SRS and WebCT quizzes for nursing students enrolled in chemistry courses showed the use of

SRS had no effect on student achievement as measured on both teacher-written and standardized exams (Bunce, VandenPlas, & Havanki, 2006). In introductory psychology courses in which questions were asked with SRS, flashcards and hand-raising, Elicker & McConnell (2011) found the most positive response from students using SRS. The response method did not, however, impact exam performance. In a study in introductory biology classes with thirteen different instructors across twelve different institutions, women were found to perform better in classes in which SRS use was integrated and men performed significantly better in the pure lecture environment. (Kang, Lundeberg, Woler, DelMas, & Herried, 2011). A similar result was found by King & Joshi (2008) in a two-semester study of engineering students in a general chemistry class. In both semesters a higher percentage of women than men were viewed to have actively participated in lecture (measure by number of SRS responses) and active women and men both scored higher on exams. A gender difference was found on exam scores, however, in that active men scored 10 points higher than non-active men whereas active women only scored 5 points higher than non-active women.

Fies & Marshall (2008) developed a framework to help guide university faculty in the use of SRS to better facilitate success. The framework, called C3, focuses on concerns, centeredness, and control. Addressing the concerns of both students and faculty about what is to be accomplished in the classroom and how to achieve these goals, clearly defining whether the goals will be achieved through teacher-centered or student-centered activities, and whether control is held solely by the instructor or shared with the students are all factors which impact the success of SRS in the university classroom.

Methods of Implementation

When it comes to student learning, the method of SRS used in the classroom (paper version vs finger indicator vs clickers) does not appear to matter in most cases. The benefit for the electronic system may be more for the instructor than the student in that electronic data can be collected and saved for instructor use at a later time, allowing greater mining of classroom data (Lasry, 2008; Price, DeLeone, & Lasry, 2010). There are some types of data which cannot be collected for analysis with a paper or hand-raising SRS such as measuring response times. In a study by Lasry, Mazur, Watkins (2008) the use of an electronic SRS allowed the researchers to delve more deeply into the nature of student responses. Results of the study demonstrated that although conceptual questions used during lecture had common detractors, students were not necessarily quick to respond to these possible answers. Student response time to questions increased when they selected a wrong answer and response time was inversely proportional to confidence about the answer.

Another advantage of an electronic SRS is the collection of demographic information about students which allows the instructor to analyze patterns in responses at a later time. These demographics can highlight underlying issues such as the impact of gender in responses. In a study by Richardson & O'Shea (2013) it was determined that although men and women in a second semester university physics course were equally likely to respond to questions correctly and in the same amount of time, men were more likely to change their initial response and to change it more often than women (students allowed to change their clicker response numerous times while a question is posed). It was also found both men and women benefited equally from peer interactions. In a study by King & Joshi (2008) engineering students' responses using clickers were tracked during a general chemistry course. The study found a higher percentage of females than males actively participating (answering 75% or more of clicker questions) in the course but active males had higher final grades (10 points) than active females (5 points) as compared to non-active members of the same gender in the course.

Feedback

The feedback mechanism, which allows both students and instructors to view immediate results on a question, is a valuable asset to the electronic SRS. In a synthesis of 134 meta-analyses of possible influences on achievement, Hattie (1982) found feedback was among the most powerful. Immediate feedback is difficult in teacher-centered, large lecture courses but easily obtained through electronic clicker systems. The value of feedback using SRS was evidenced in a study by Doucet, Vrins, and Harvey (2009) where nearly 100% of students "indicated that the ability to evaluate their own strengths and weaknesses was the single most useful advantage of using the ARS [sic... audience response system]." In a study on the effectiveness of peer instruction with SRS in an engineering dynamics course, Schmidt (2011) demonstrated the use of clickers provided a significantly valuable means of self-assessment of the academic outcome in the course. In a multi-method study across five course in three science disciplines, Hoekstra & Mollborn (2012) demonstrated SRS can be used to support existing pedagogical goals through teaching five practices – gathering student feedback to improve teaching and learning, identifying student assumptions or preconceptions about course material, supporting conceptual application and critical thinking through small- and large-group discussions, fostering social cohesion in the learning community, and collecting data from students to support theory testing, conceptual application, and group discussion.

Nature of Discussions

Many classrooms using SRS now include the use of peer instruction, which incorporates student-student (or small group) discussion about the questions asked in class (Mazur, 1997). The timing and nature of this discussion can vary. Some instructors have students respond individually first, then discuss their answers with one or more other students before replying to the question a second time (re-polling). Other instructors will have students immediately discuss the question with others before having students answer the first time (Dufresne, 1996). Discussion with classmates has been shown to increase the percent of students ultimately selecting the correct answer on the question, especially when the initial percentage of correct answers is around 50% (Crouch and Mazur, 2001). Peer instruction, which includes questions focusing on student misconceptions, has been found to improve conceptual and quantitative understanding of physics material in both introductory calculus and algebra-based physics courses (Crouch and Mazur, 2001).

Clearly something of value occurs during the discussion. Research studies within the past several years have begun to look at the nature of the discourse occurring during the peer learning interaction. Smith, Wood, Adams, et.al (2009) found discussion enhances understanding even if no one in the group knew the correct answer. James and Willoughby (2011) recorded student conversations during the peer interaction and discovered three broad categories of "non-standard" (not focused on the specific question detractors or question concepts) conversations: 1) student ideas about the subject that were unanticipated by the question answers; 2) statistical feedback misrepresenting student understanding such as extraneous cues or the confidence of the partner; and, 3) conversation pitfalls related to students ability to engage in conversation effectively. Nicol & Boyle (2003) found the type of dialogue and the discussion sequence used in conjunction with the clickers had important effects on learning. Students using peer instruction with one or two classmates was more beneficial to learning than class-wide discussion. Larger discussions led to greater student confusion and decreased attention and motivation. Webb and Mastergeorge (2003) found certain student behaviors were necessary for effective help-seeking and help-giving which occurs between discussion partners in peer instruction. Help seekers need to be precise in asking questions, persist in seeking help, and apply the explanations they received whereas effective help-givers need to provide detailed explanations and opportunities for those they helped to apply what they discussed.

James (2006) found peer discussions were impacted by grading incentive. In this study it was found that if astronomy students were given little or no credit for incorrect responses, the peer partner with greater knowledge dominated the conversation. If the penalty for incorrect responses was removed, there was more equal distribution of discussion between the partners. Willoughby & Gustafson (2009) found similar but inconsistent results in their classrooms. During the first semester of their study comparing high and low stakes use of clickers in two astronomy course sections, there were statistically significant differences found in the average number of correct responses and the degree of block voting (providing one four-member group answer) with students in high stakes classroom more likely to select the correct answer and block vote. These differences were not seen in the second semester of the study. In neither semester was there a statistically significant difference in average gain on an Astronomy Diagnostic Test or in course final grade. In an analysis of recorded discussions during clicker questions, it was found students in low stakes classrooms were much more likely to restate the questions, state answer preferences and ask for question clarification.

The value of conceptual questions with clickers and peer discussion can be enhanced with instructor explanation. In a study by Smith, Wood, Krauter and Knight (2011) clicker questions were presented with three formats - peer discussion only, listening to instructor explanations only, or with peer discussion followed by instructor explanation. Results showed peer instruction followed by instructor explanation did more to improve overall student performance than either of the other two formats. Slight differences were found based on ability level. Weak performing non-majors show somewhat more benefit from instructor explanation alone whereas the same technique showed no effect with top-performing students.

If the conversation impacts the nature of the response then optimizing that interaction is of value. Since student ideas about the subject, extraneous cues, and confidence level appear to matter in the student-student discourse, a student's previous experience with the subject matter may be a factor in the success of peer instruction. In a large national study of introductory college students (Rudolph, et.al., 2010) the connection between numerous student demographics and student learning with interactive learning strategies were measured. Demographics in this study included such things as gender, previous astronomy class, previous math and physical science courses, major, year in college and many others – 15 demographics in total. While it was found that year in school, GPA, highest level of math and number of physical science courses taken did have an impact on student improvement in the interactive classroom, students who had taken a previous astronomy class did not outperform their fellow students. This lack of impact of previous subject experience on college success has not been demonstrated in all science disciplines.

In a study of 1,933 introductory college physics students, Sadler and Tai (2001) found, controlling for student backgrounds, students who took a high school physics course were more likely to receive a higher grade in introductory college physics than students with no previous subject experience. In a study of 8,310 introductory college biology, chemistry, and physics students in 55 randomly selected U.S. colleges and universities, Schwartz, Sadler, Sonnert & Tai (2008) found students who had spent more than a month covering a particular topic in high school science earned higher grades in college science than those students who reported no in-depth coverage of material. Tai, Ward, and Sadler (2006) found students who had substantial coverage of stoichiometry in high school were more successful in introductory college chemistry courses.

Summary

The current mixed methods study focuses on the influence of previous subject experience on the level of student success and interaction during peer instruction in an introductory, one-semester, college-level Survey of Physics course. While a substantial body of research, both qualitative and quantitative, exists on the use of peer instruction and student response systems, little consideration has been given as to the impact of previous subject experience in the nature of student responses and interaction. Most research studies have also been purely qualitative or quantitative, lacking a mixed-methods approach to provide a more comprehensive interpretation of the student experience during peer instruction.

CHAPTER 3: RESEARCH DESIGN AND METHODOLOGY

Mixed Methods Research Design

This study uses a mixed methods research design, a research methodology in which quantitative and qualitative methods are used in combination with the intent of providing a better understanding of research problems than either method alone (Creswell & Plano Clark, 2011). The mixed methods design incorporates the statistical, closed-end research design found in quantitative research with the storied, more personal, open-ended aspects of qualitative research. This combination of data and analysis can "simultaneously address a range of questions through both qualitative (exploratory questions and theory generation) and quantitative (confirmatory questions and theory verification) approaches, provide stronger inferences, and provide for more divergent views." (Teddlie & Tashakkori, 2009, p. 33).

Quantitative, qualitative, and mixed methods describe different research methodologies. A research methodology is a "broad approach to scientific inquiry specifying how research questions should be asked and answered" (Teddlie and Tashakkori, 2009, p. 21). Mixed methods research incorporates methodologies of quantitative and qualitative research but also stands on its own as a methodology since both research methods are combined and analyzed. Mixed methods research has been called the "third research paradigm" (Johnson & Onwuegbuzie, 2004, p. 15), following the development of quantitative and qualitative research methodologies.

According to Greene (2006), any methodology should have four domains:

1. Philosophical assumptions (paradigm)

2. Inquiry logics – inquiry questions and purposes, broad inquiry designs and strategies, sampling logic, criteria of quality, etc.

3. Guidelines for practices - specific methods for conducting inquiries such as sampling strategies, analysis techniques, etc.

4. Sociopolitical commitments in science, concerned with values like "whose interest served is by particular approach to inquiry" (paradigm)

Philosophical assumptions. Tashakjori and Teddlie (2003) describe pragmatism as the primary paradigm of mixed methods research. In this paradigm both quantitative and qualitative methods may be used in a single study. The research questions are of primary importance, taking priority over method or philosophical worldview. A pragmatic paradigm does not include a forced choice between post-positivism and constructivism and abandons the concepts of "truth" and "reality". The focus of the pragmatist paradigm is that a practical and applied research philosophy should guide methodological choices (Teddlie & Tashakkori, 2009). Not all mixed methods researchers believe in the pragmatism world view. Greene and Caracelli, (2003) believe multiple paradigms can be used in mixed methods studies and Creswell, et. al. (2003) describes six different mixed methods designs and how paradigms differ between methods.

Explanatory Sequential Design. This study utilizes an explanatory sequential mixed methods design, a two-phased design in which qualitative data helps explain or build upon initial quantitative results (Creswell & Plano Clark, 2011). The quantitative data may be significant or non-significant, have outliers or surprising results (Morse, 1991) or be used to form groups for follow-up qualitative study (Teddlie and Tashakkori, 2009).

The two-phased structure of the explanatory design makes it straightforward to implement since there are two different methods undertaken at different times. The design lends itself to multiphase investigations as well as single mixed methods studies and is generally favored by quantitative researchers since it begins with a strong quantitative study (Teddlie & Tashakkori, 2009). The general format of a sequential explanatory design is shown in Figure 1 (Creswell & Plano Clark, 2007).

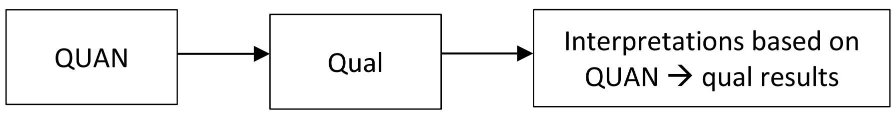


Figure 1. Explanatory sequential mixed methods design structure. Emphasis of the quantitative component of the study is indicated by using all capital letters.

For this study the mixed methods approach will utilize the “Follow-up Explanations” sub-model of the explanatory design. (Creswell & Plano Clark, 2011). A diagram of this method is provided in Figure 2 and includes a summary of procedures and products for the various components of the study.

There are challenges associated with all types of mixed methods research including the explanatory sequential design. In addition to managing a lengthy research project, the researcher must also consider how to deal with varying sampling techniques and sizes in both aspects of the study as well as how to integrate the quantitative and qualitative data (Teddlie & Tashakkori, 2009). These challenges are addressed in the data collection and analysis sections.

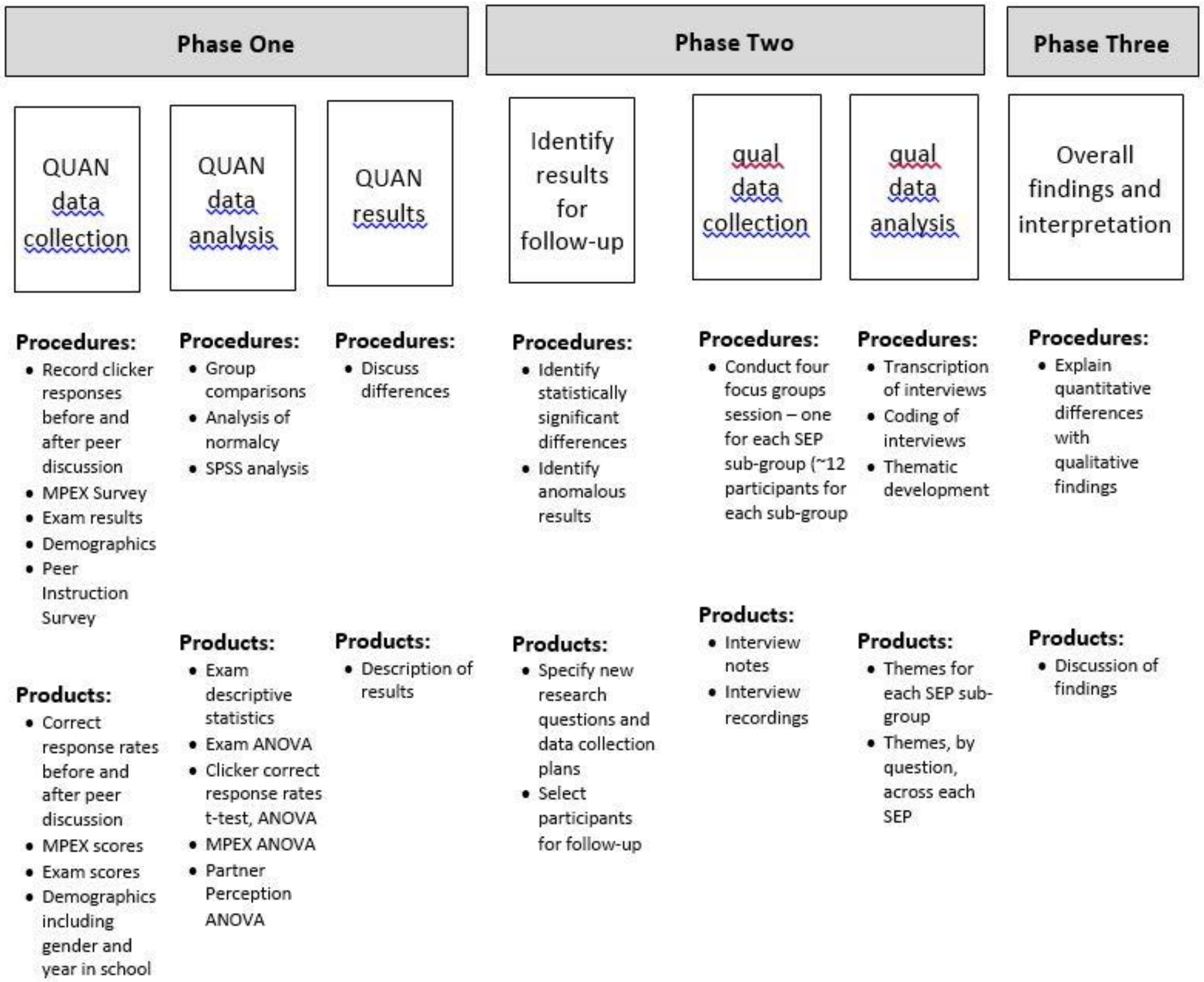
Target Population and Sampling

Due to the requirement of the use of clickers in the classroom (which has a financial requirement on the part of the student and a use commitment on the part of the instructor) and the specific peer instruction technique desired, there were limited options for choosing a physics class and instructor for this study within this researcher's geographic area. Therefore, the population for this study was a sample of convenience – the researcher's one-semester "Survey of Physics" course. This course is an introductory, freshmen-level course offered at a four-year university in the northern plains. Typical enrollment each semester is approximately 120 students. This course meets the general education science requirement at the institution but is also a required science class for nearly all agricultural majors on campus. The Survey of Physics course is comprised of three hours of lecture, a two-hour lab, and a one-hour recitation each week. Data for this study was collected during the lecture and lab components with some data collected in an online survey.

Data Collection and Analysis

This study was conducted in three phases. Phase 1 focused on quantitative data collection and analyses including clicker responses before and after peer discussion, results from the Maryland Physics Expectations (MPEX) survey, unit exam scores, and results from a peer instruction partner survey. Phase 2 addressed the qualitative aspect of the study with purposeful sampling of each SEP membership (P of PP, P of PN, N of PN, and N of NN) for participation in semi-structured interviews. In Phase 3 the mixed

Figure 2. Detailed overview of Follow-up Explanations form of Explanatory Sequential Design utilized in this study. Modified from Creswell & Plano Clark, 2007.



research question was addressed through integration and synthesis of findings from Phases 1 and 2.

Phase 1 - Quantitative Data Collection: Demographics and Partner Assignments. During the first week of class, students completed the MPEX survey as part of the required physics department assessment plan. Three additional demographic questions were added to the MPEX asking students their year in school, gender and whether or not they took a high school physics course. Utilizing the MPEX demographic results, students were considered a member of either the group who had high school physics (designated ‘P’) or the group who did not have high school physics (designated ‘N’). A stratified sampling technique was then be used to assign each student a discussion partner for peer instruction resulting in three subject experience pairs (SEP) – PP, PN, and NN. The demographics of gender and year in school were distributed, as equally as possible, between each SEP to remove these characteristics as confounding variables. During the ninth week of the semester, students were paired into one of the three SEP and assigned seating during the lecture component of the course.

Phase 1 - Quantitative Data Collection: Student response systems. The SRS is a handheld device that allows students to electronically input responses to multiple-choice questions posed on a PowerPoint slide. For this study the instructor asked questions and collected student responses utilizing the Turning Technologies SRS which included software embedded with PowerPoint and a radio receiver. Students in this study were required to have an SRS as part of the course and each clicker was registered so electronic responses could be linked to each student. The instructor for the Physics 101 course has had over 10 years of experience with clickers prior to this study and has given presentations and training sessions on their use.

During the first part of the semester, SRS questions were infused into the lecture and students were required to enter responses with their handheld device. Students had not yet been assigned into a SEP and peer instruction was not yet be part of the course. The purpose of this initial introduction of SRS was to provide students a learning period of how to use the clicker, including setting the proper channel, how to interpret if their responses had been recorded and the timing and nature of questions in lecture.

During the ninth week of the semester, students were required to sit in their assigned seats with their assigned discussion partner. SRS and peer instruction were used in the classroom for a week but the data was not part of this study. The intent of this week was to acclimate students to the process of peer instruction in the classroom. This included the sequencing (answer alone, discussion, answer again) of questions and the expectations for discussion with their assigned partner. Allowing a time for students to become comfortable with each other for the purposes of peer instruction is implied by Uncertainty Reduction Theory (Berger & Calabrese, 1975) which posits people have a need to reduce uncertainty about others by gaining information about them. This can either be done through passive means such as general observations or active strategies such as discussions with the other person. To alleviate the uncertainty of working with a stranger for peer instruction, time was be allotted in the first class in which SEP partners were assigned, for students to introduce and learn something about each other. Students also gained familiarity with each other during this first week of peer instruction as they discussed answers to SRS questions posed in lecture.

This first week of using SRS was intended to remove any effects of varying experience with clickers or peer instruction among the SEP. After this introductory period, quantitative data collection occurred over the remainder of the semester, a period of seven weeks.

For each clicker question asked during lecture (and posted on a PowerPoint slide), students first answered the question by themselves. After their initial response was entered, students were asked to discuss their answer with their partner and try to resolve any differences. The same question was re-polled with all students entering an answer the second time. After the second entry, a chart showing the response distributions was displayed on a PowerPoint slide and the correct answer revealed to the class as a whole. If less than 70% of the class answers correctly, the instructor provided a more in-depth explanation of the answer while soliciting input from the class. If more than 70% of students had the answer correct, a more limited explanation of the correct answer was provided. A comparison slide of correct-incorrect responses on the first and second attempt on each question was provided to the class with the intent of helping demonstrate the value of discussion during the peer instruction process.

Clicker data from each lecture was saved using the TurningPoint software and later imported into a Microsoft Excel spreadsheet in which students were grouped by partner and category (P of PP, P of PN, N of PN, or N of NN). The spreadsheet was used to calculate correct response rates on questions prior to and after partner discussions. Data was then be uploaded into SPSS and coded for analysis, removing student names.

Phase 1 - Quantitative Data Collection: Peer instruction and academic achievement. To evaluate whether peer instruction within a SEP had an impact on student learning outside the peer instruction environment, exam scores for each peer instruction pair were compared and statistically analyzed using SPSS. During the quantitative phase there were two unit exams.

Phase 1 - Quantitative Data Collection: Student attitudes and perception. Student perceptions about peer instruction and learning physics was measured two ways. At the end of the study, students were asked to complete six Likert-scale questions related to their perceptions about their discussion partner during peer instruction. The survey was completed within the online survey tool, QuestionPro. Data was downloaded from QuestionPro into a Microsoft Excel spreadsheet in which students are grouped by partner and SEP membership (P of PP, P of PN, N of PN, and N of NN). Since there are only six Likert items, student ratings on each item were combined to develop one overall Peer Instruction Partner Score (PIPS) for each student. The data was then uploaded to SPSS and coded for analysis, removing student names.

Student changes in perceptions about the nature of physics and how to be successful in studying physics were measured through pre/post responses to the MPEX. This assessment was a normal part of the physics department assessment plan for this course. During the first week of class, the MPEX was provided electronically to students using the online survey software, QuestionPro. The MPEX survey consists of thirty-four statements about the nature of physics and how students think they would best learn physics. Each statement is a Likert item which students rank from 1-5 based on their level of agreement. Students took the MPEX again near the end of the semester, allowing for a comparison of pre/post data and the ability to evaluate the influence of SEP membership (P of PP, P of PN, N of PN, N of NN) on shifts in attitudes.

Students were required to enter their name on the MPEX survey so that answers could be associated with a specific student. Data was downloaded from QuestionPro into a Microsoft Excel spreadsheet in which students were grouped by partner and SEP category. The data was then uploaded to SPSS and coded for analysis, removing student names.

Phase 1 - Data Analysis: SRS, Previous Subject Experience, and Peer Instruction. All clicker responses for each student were uploaded into a Microsoft Excel spreadsheet for analysis. The number of correct responses prior to and after the partner discussion were totaled for all clicker questions asked during the period of this study. The analysis began with an overall comparison of first correct response rates - participant responses prior to peer discussion. An independent samples t-test was used to determine if a difference existed in correct response rates between students who had high school physics (P group) and those who did not have high school physics (N group). The independent variable was previous subject experience and the dependent variable was correct response rate.

To determine the influence of SEP membership – P of PP, P of PN, N of PN, and N of NN – on correct response rates, the correct response rates after peer discussion were compared. A one-way ANOVA was used to analyze correct response rates and SEP membership (P of PP, P of PN, N of PN, and N of NN). The independent variable was SEP membership and the dependent variable was correct response rate.

Phase 1 - Data Analysis: SEP membership and academic achievement. For the two unit exams taken during the time of this study, descriptive statistics and the ANOVA statistic was applied using SPSS. The independent variable was SEP membership (P of

PP, P of PN, N of PN, and N of NN) and the dependent variable was exam score. The statistic was used to determine if there was a statistically significant difference between the means of the exam scores and membership in a SEP.

Phase 1 - Data Analysis: SEP membership and student perception. Student perceptions about their partner during peer instruction was analyzed on the Peer Instruction Partner Survey (Appendix C). For each student the data collected on the six-question Likert survey was combined to obtain one average Peer Instruction Partner Score (PIPS). The mean score for each SEP membership was analyzed using an ANOVA with the independent variable being SEP membership (P of PP, P of PN, N of PN, and N of NN) and the dependent variable PIPS score.

For each student, results from the 34-item MPEX survey was compiled to obtain a percent favorable score. Favorable is defined as a rating of either a 4 or 5 or a 1 or 2 on each Likert item, depending on whether or not experts agreed with the statement. A response of three was considered neutral. According to the meta-analysis of the MPEX literature, the most common means of evaluating the results of the MPEX involved the analysis of the percentage agreement with experts and the shift in this agreement over time (Madsen, et.al., 2015; Elby, 2001). For this study an average percent agreement score was obtained for each SEP membership – P of PP, P of PN, N of PN, and N of NN. The mean percent agreement scores for each SEP membership was statistically compared using a repeated measures ANOVA, comparing scores at the beginning of the study to those at the end of the study. In this analysis, SEP membership was the independent variable and the mean percentage agreement score was the dependent variable.

MPEX survey items were aggregated into six dimensions - independence, coherence, concepts, reality link, math link, effort (Table 1), and evaluated as to their level of favorability (agreement with experts). For this study a repeated measures ANOVA statistic was applied using SPSS to evaluate the significance of changes in percent agreement for each dimension, where appropriate. As with the overall MPEX score, the repeated measures ANOVA compared pre/post SEP membership results. For each

Table 1

Rating Descriptors and Item Grouping for MPEX Survey

Dimension	Favorable	Unfavorable	MPEX Items
independence	takes responsibility for constructing own understanding	takes what is given by authorities (teacher, text) without evaluation	1, 8, 13, 14, 17, 27
coherence	believes physics needs to be considered as a connected, consistent framework	believes physics can be treated as unrelated facts or "pieces"	12, 15, 16, 21, 29
concepts	stresses understanding of the underlying ideas and concepts	focuses on memorizing and using formulas	4, 19, 26, 27, 32
reality link	believes ideas learned in physics are relevant and useful in a wide variety of real contexts	believes ideas learned in physics has little relation to experiences outside the classroom	10, 18, 22, 25
math link	considers mathematics as a convenient way of representing physical phenomena	views the physics and the math as independent with little relationship between them	2, 6, 8, 16, 20
effort	makes the effort to use information available to tries to make sense of it	does not attempt to use available information effectively	3, 6, 7, 24, 31

ANOVA the independent variable was SEP membership and the dependent variable was percent expert agreement.

Phase 1 – Summary. The focus of the quantitative portion of the study was to determine the influence of previous subject experience in four ways. The first analysis focused on correct response rates to clicker questions when students respond independently, prior to peer discussion. This established the influence of previous subject experience in initial clicker responses. The second analysis focused on correct response rates to clicker questions after peer discussion. This analyzed the influence of the discussion partner in terms of previous subject experience (P of PP, P of PN, N of PN, and N of NN). The third analysis considered the extent of the influence of the SEP membership on learning by assessing the results of exam scores for each membership. The final analysis considered the extent of the influence of the SEP membership in terms of perceptions about peer instruction and physics. This was analyzed by the Peer Instruction Partner Survey and the MPEX. A summary of the quantitative analyses used in this study is provided in Table. 2.

Table 2

Quantitative Data Collection and Analysis Summary Table

Research Question	Assessment	Statistic	Independent Variable	Dependent Variable
What is the influence of previous subject experience on SRS correct response rates prior to peer discussion?	Correct response rates on clicker questions posed in lecture – prior to peer discussion	t-test	Previous subject experience: Designation of 'P' for students who took high school physics and a designation of 'N' for students who did not take high school physics	Mean correct response rates for P and N
What is the relationship between previous subject experience and SRS correct response rates after peer discussion for each SEP?	Correct response rates on clicker questions posed in lecture – after peer discussion	One-way ANOVA	SEP Membership: P of PP P of PN N of PN N of NN	Mean correct response rates for each membership
What is the influence of previous subject experience and peer instruction on student understanding of the subject matter as measured on unit exams?	Two unit exams	Descriptive Statistics and One-way ANOVA	SEP Membership: P of PP P of PN N of PN N of NN	Mean exam scores for each membership
What is the influence of previous subject experience and peer instruction on student perceptions about their discussion partner, the nature of physics and how to be successful in a physics course?	Partner Perception Survey (PSS)	One-way ANOVA	SEP Membership: P of PP P of PN N of PN N of NN	Mean PSS score for each membership
	Maryland Physics Expectation Survey (MPEX)	Repeated measures ANOVA	SEP Membership: P of PP P of PN N of PN N of NN	Mean percent expert agreement for each membership

***Phase 2 - Data Collection: Focus groups*.** The qualitative component of the study was conducted at the end of the quantitative component, after about a week of completing the quantitative analysis (which was continuous during the last seven weeks of the semester). During the qualitative component, students were purposefully sampled from each of the SEP (P of PP, P of PN, N of PN and N of NN) for selection into focus groups for semi-structured interviews. Focus groups provide "a way of collecting qualitative data, which – essentially – involves engaging a small number of people in an informal group (or discussions), 'focused' around a particular topic or set of issues" (Wilkinson, 2004, p. 177). Focus groups were chosen over individual interviews because past research has shown participants in focus groups feel more comfortable in the socially-oriented environment (Krueger, 2000) and feel safe to share information (Vaughn, Schumm, & Sinagub, 1996). In this study the semi-structured interviews in each focus group were designed to gather information to explain the results from the quantitative phase of the study, in particular to more fully understand student perceptions and experiences during peer instruction.

Focus groups should have between six to twelve participants (Krueger, 2000; Johnson & Christensen, 2004). In this study, every effort was made to maximize student participation in each focus group to provide the greatest possible input from the sampled SEP. An equal mixture of gender and year in school was also purposefully selected for each of the focus groups to remove potential bias in the sample. Each student selected for the focus group was sent a personal invitation from the instructor via email to participate in the focus group. The invitation clarified that student participation in the focus group was voluntary and that although recorded, no comments would be attributed to any particular individual. To incentivize student attendance in each focus group, students were provided pizza and pop prior to the discussion and a $10 Wal-Mart gift card upon completion of their participation in the focus group. Students were also given the option to enter their name into an overall drawing for a $50 gift card for each session.

During each focus group, a semi-structured interview was conducted. Questions during the interview included four types: essential, extra, throw-away, and probing questions (Berg, 1998). Throw-away questions were asked at the beginning of the interview and were designed to put participants at ease and start the conversation. These included asking participants to introduce themselves and discuss their previous experience with SRS. The essential questions focused on the primary purpose of the study and were prepared in advance as were some of the extra questions which tried to come at the topic from a slightly different perspective. Both the essential and extra questions for this study focused on students' perceptions and experiences with using SRS and peer instruction in the physics classroom. Probing questions were used to get participants to expand on a particular topic and were not developed in advance. The essential and extra questions designed for the semi-structured interview focused on:

- Previous experience with SRS
- Perceptions about the value of SRS in learning physics
- Perceptions about the value of peer interactions
- Explanations about the nature of discourse between subject experience pairs during peer interactions
- Shared experiences about challenges and successes in learning physics

The full interview protocol form, including essential and extra questions is included in Appendix B.

Each focus group session was recorded. The tapes for each focus group were transcribed by the researcher and analyzed for codes and thematic development.

Phase 2 - Data Analysis: Focus groups. In qualitative research "the vast array of words, sentences, paragraphs, and pages have to be reduced to what is of most significance and interest" (Seidman, 2013). Analysis of the data began with the electronic copies of the transcripts from each focus group. Only one transcript, one SEP, was evaluated at a time. Each transcript was read several times before coding began and continued through the coding process. The goals of this iterative process was to find patterns within that data that related to the research questions. Patterns within qualitative data can have many forms. Hatch (2002) suggests characterizing patterns based on similarity, difference, frequency, sequence, correspondence, and causation. Once coding was completed for a given question, broader themes were developed as they emerged, within and across each SEP. The process was repeated for each SEP.

Validity and Reliability

For a research study to be of value, questions of validity and reliability must be addressed. The approach to validity and reliability varies between quantitative and qualitative research. These topics are discussed separately for each phase of this study.

Phase 1: Quantitative validity and reliability**.** Although this study did not utilize a fully random sample design since the sample was drawn from the researcher's course, it did have multiple measures with the instrumentation used which supports its quasi-experimental design nature (Trochim, M.K., 2006). The NN group serves as the comparison group to which the impact of a student with previous physics experience in a group was being compared. Within quasi-experimental design studies there are more threats to internal validity than in a true experimental design. Single group threats can include selection-maturation, selection-history, selection-regression, selection-testing, selection-instrumentation, or selection-mortality (Trochim, M.K., 2006). Selection-maturation can lead to an interpretation of a cause-effect relationship when one does not exist because changes in pretest and posttest score are really the result of groups maturing at different rates. Selection-history clouds causal results because the results may be due to how a particular group within the study reacted differently to some event which occurred or may be due to one group experiencing an event the other one did not. Selection-regression is seen when group improvement is really just a regression to the mean. Selection-testing can confound results when different groups experience different priming effects on the posttest as a result of taking the pretest. Selection-instrumentation relates to any change in the test used for the groups between pretest and posttest, including the method in which the test is given or method of evaluation. Selection-mortality confounds results when there is a differential, nonrandom dropout between pretest and posttest (Trochim, M.K., 2006).

To strengthen this quasi-experimental research design, relevant threats to the internal validity were considered. Selection-testing and selection-instrumentation threats are considered least likely to impact the study. Selection-testing does not apply in this case since the MPEX survey was administered over the semester and is only attitudinal in nature. Selection-instrumentation was not considered to be a threat to validity since the MPEX was administered in the same way in both the pre- and post-assessments and the interpretation of the results is quantitative and not subject to this researcher's changing ability to interpret the results.

Selection-history was minimized since all experience pairs (PP, PN, NN) were in the same course section with all pairs exposed to the same instruction in the same manner from the same instructor. The stratified sampling used for assignment into experience pairs also weakens the potential of this threat. Data collected to consider the impact of selection-history is demographic in nature (gender, year in school, randomness of experience pair). If significant differences in demographics exist between the experience groups, there would be a greater likelihood of a selection-history threat. Selection-regression was analyzed by looking at the pre/post results for similar improvement or change. If, for example, all experience pairs start below the mean and improve toward the mean, the result may be due to selection-regression and not whether or not a pair contained a person with previous physics experience. The threat of selection-maturation was evaluated by again considering the pre/post test results and to what extent the scores changed over time. Selection-mortality was analyzed by keeping track of the responses of the experience pairs throughout the process. Some students within a pair had dropped the course and others did not attend class on a regular basis. Substantial differences in selection mortality between the experience pairs can impact the internal validity of the study.

Instrumentation. To assess attitudinal changes during the portion of the course in which data was collected, the Maryland Physics Expectation Survey (MPEX) was administered to students. The MPEX survey was developed at the University of Washington in the fall of 1992 (Redish, Saul, Steinburg, 2000) as a means of assessing students' cognitive expectations of a physics course as related to their understanding of the process of learning physics and the structure of physics knowledge. The design of the MPEX survey had three primary goals: 1) to determine how the initial state of students in university physics differed from the view of experts; 2) to understand the extent to which the initial state of a class varied from institution to institution; and, 3) to understand how the expectations of a class changed as a result of one semester of instruction in various learning environments. The authors of the study posited the value of measuring expectations about a physics course since a students' expectations affect how they respond to a course. In particular student expectations impact how they filter information in the course and which activities they select to help them construct their knowledge. The authors felt there was strong potential impact in situations where "there is a large gap between what the students expect to do and what the instructor expects them to do" (Redish, et. al., 2000).

The MPEX survey uses a five-point Likert scale to measure students' attitudes, beliefs, and expectations about the introductory physics course. The survey looks at six specific dimensions - independence, coherence, concepts, reality link, math link, and effort. Independence considers whether the student learns independently or simply takes in information from the instructor without evaluation. Coherence considers whether the student sees physics as a connected framework or rather as a set of individual facts. Concepts considers whether students focus on and value the learning of the concepts or whether they focus more on just the mathematics - memorizing and using formulas. Reality link evaluates to what extent students believe the material covered in physics is relevant and useful in the everyday world. Math link relates to how well a student sees the connections between the math and physics. Effort considers the extent to which students use the information provided and tries to make sense of it. Developers of the MPEX survey have developed categories of "favorable" and "unfavorable" for each of the dimensions described above. The "favorable" category aligns with a view of most "expert" physics instructors with a high concern for educational issues and who were deemed sensitive to students. The "unfavorable" category reflects a novice view common to most beginning students and was defined as being in disagreement with the expert responses.

Items for the MPEX survey were developed from data obtained through literature reviews, discussions with physics faculty and the decades of experience of the survey authors. The content validity of the survey was tested at numerous universities and colleges over the course of four years. The survey authors conducted more than 100 hours of videotaped student interviews to validate that students interpreted the survey items as they were intended by the authors. In a more recent study Omasits and Wagner (2006) interviewed students about their responses on the MPEX items and found 95% of student explanations were consistent with the corresponding Likert choice.

In the development of the MPEX survey, the instrument was given to five calibration groups consisting of engineering students in a calculus-based physics course, members of the U.S. International Physics Olympics Team, high school teachers attending a two-week seminar on new approaches to physics education, university and college faculty attending the same summer seminar, and college faculty who were part of the Workshop Physics project. On all but three items, the Workshop Physics faculty had 87% agreement on the 34 Likert items. This was established as the "expert response". The three items not receiving 87% agreement (items 7, 9, and 34) still had strong agreement but the neutral choice was selected by 25-33% of the respondents. Of the remaining calibration groups, the university/college faculty attending the summer workshop had 80% agreement with the expert response, high school physics teachers had about 73% expert response agreement, the Olympiad team a 68% agreement and the college physics students a 54% percent expert agreement.

The MPEX was then given to students in calculus-based physics courses at six different institutions (N~1500). Three institutions were Tier 1 research institutions, two were liberal arts colleges and one was a medium-sized, two-year college. All courses were covering the topic of Newtonian mechanics but the method of instruction varied. There was continued consistency in the expert response by instructors at these institutions, providing further support to the “expert response” category. On the MPEX pretest, students consistently had substantially different views than physics experts (50-60% favorability) and rather than indicating a neutral response, students chose unfavorable (disagreement with experts) 15-30% of the time. The concept cluster only had a 30-45% agreement with expert response whereas the reality cluster had a 60-75% agreement with expert response. The results from students at each of the Tier-1 research institutions were similar whereas the responses from students at the liberal arts schools were consistently on the higher end and results from the two-year college consistently on the lower end.

On the MPEX student post-tests, the majority of percent favorable responses either remained unchanged or decreased, regardless of the method of instruction or size of school. The largest deteriorations were in the dimensions of effort, coherence and reality link. A statistical analysis of the results led the authors to conclude that a shift in means of 5% was significant for large schools (N>=450) and a shift in means of 10% to be significant for smaller schools.

The MPEX has served as a measure of student expectations in physics education research for decades. The existence of multiple studies has provided the opportunity for a more in-depth evaluation of attitudinal survey results in physics classes. In a meta-analysis of 24 studies using the Colorado Learning Attitudes about Science Survey (CLASS) and the MPEX, the correlation between student perceptions and external factors has been clarified (Madsen, McKagan, and Sayre, 2015). In a one-way ANOVA with teaching method as the independent variable and shift as the dependent variable, a significant main effect was found ($F(3,52 = 698.2$, $p<.001$). In courses where there is an explicit focus on developing student expert-like beliefs, overall positive gains of 8.5% have been measured. Large positive gains in expert-like responses also result from courses focused on developing models of the physical world (9.3% shifts). In modeling courses students work in small groups to perform experiments and gather evidence to build models.

Class size also appears to be a factor in CLASS/MPEX score improvement. In a one-way ANOVA with class size as the independent variable and shift as the dependent variable, a significant main effect was found ($F(2,53)=5.4$, $p=.007$). In the meta-analysis more positive shifts to expert-like responses occurs in smaller classes (5.4%) whereas large classes tend to result in a negative shift in expert-like beliefs (-1.7%).

The third factor influencing shifts in CLASS/MPEX scores is the nature of the class population. A one-way ANOVA with the independent variable as population and shift as the dependent variable show a significant main effect for student population ($(F3,52)=5.8$, $p=.002$). Large positive shifts in perception have been measured in physics courses for elementary education and non-science majors (6.7%) but no change or negative shifts occur in algebra and calculus-based physics courses (-2.8 to 0.2%).

Phase 2: Qualitative validity and reliability

Healy and Perry (2000) suggest the quality of a study should be judged by the terms of a research method's paradigm. While the terms validity and reliability were historically developed around the quantitative research paradigm, they do not transfer well to the qualitative research paradigm. Lincoln and Guba (1995) suggest the trustworthiness and worth of qualitative research depends instead on the establishment of credibility, transferability, dependability and confirmability. This study will focus on credibility and dependability which are the closest in likeness to validity and reliability.

Credibility describes the confidence in the truth of the findings which is broadly similar to the concept of validity in quantitative research. Dependability involves the demonstration that the findings are consistent and could be repeated (Lincoln & Guba, 1995). In this study credibility and dependability will be established through member-checking and triangulation. Member-checking involves providing the findings to study participants and asking them whether the findings are an accurate reflection of their experience (Creswell & Plano Clark, 2011). In this study member-checking involved providing participants of each SEP focus group a summary of the major themes developed from the interview, allowing participants the opportunity to correct errors and challenge interpretations of the data.

Triangulation involves drawing data from several sources or from several individuals (Creswell & Plano Clark, 2011). In the qualitative part of this study, triangulation is achieved through building evidence of codes from several different individuals.

Phase 3: Mixed methods validity and reliability

The strength of mixed methods research is in the collection of a broad range of data, both qualitative and quantitative. This broad spectrum of data can be used to validate findings through triangulation, in particular using methods triangulation (Denzin, 1978; Patton, 1999). This was accomplished by comparing the themes which develop from the qualitative study with the results of the quantitative study.

Ethical Consideration

Permission to conduct this study was obtained from both the researcher's institution (SDSU IRB Approval # IRB-1609006-EXM) and from the IRB at the University of Nebraska (IRB# 20161116648 EX). Although it was necessary to collect student names to correlate clicker numbers with respondents and associate survey results with members of each SEP, all data was coded to provide anonymity for each participant and names were kept confidential. All data was kept password protected on the researcher's university-issued computer in her locked office. Any hard-copies of data were kept in a locked file cabinet in the researcher's office.

Summary

This chapter provided an overview of the research methodologies, participants, and data analysis for this study. A summary of the mixed methods design, measures and products from each part of the study is provided in Figure 2. Findings are presented in the next chapter and correlated with each research question.

Chapter 4

Quantitative and Qualitative Results

This chapter provides the results of the data analyses described in the previous chapter. The quantitative results are presented by research question and the qualitative results by theme and SEP.

Phase I – Quantitative Results

Descriptive statistics of participants. Participants in this study included students in a one-semester Survey of Physics course at a four-year university in the Northern Plains. Although the study began with 123 students, students who dropped before the completion of the study (*n* = 3) or had class attendance in excess of 1.5 standard deviations from the average (less than 52.6% attendance) were excluded from the study. There were 113 eligible participants for the study including 36.3% females (*n* = 41) and 63.7% (*n* = 72) males. The class distribution included 2.7% freshmen, 22.1% sophomores, 46.0% juniors, and 29.2% seniors.

Research question 1 analysis. An independent samples *t*-test was used to examine the correct first-response rates of the P students (M = 64.97%, SD = 11.48%) and N students (M = 57.44%, SD = 13.12%). The hypothesis was students who had taken a high school physics course (P group) were more likely to answer clicker questions correctly prior to peer discussion than those who had not taken high school physics (N group). As noted in Table 4.1 the distributions of the P and N groups were sufficiently normal for the purpose of conducting a *t*-test (i.e., skew$<|2.0|$ and kurtosis$<|2.0|$; Trochim & Donnelly, 2006; Field, 2009; Gravetter & Wallnau, 2014). In addition the assumption of homogeneity of variance was tested and satisfied according to Levene's *F* test, $F(111) = 1.307$, $p = .255$.

Table 4.1

Descriptive Statistics Associated with Clicker First Response Rates

Group	N	M	SD	Skewness	Kurtosis
P	50	64.97	11.48	-.553	.938
N	63	57.44	13.12	-.265	-.154

The independent samples *t*-test indicated a statistically significant effect, $t(111) = 3.200$, $p = .002$. These results suggest students with previous subject experience are more likely to answer a clicker question correctly before peer discussion. The Cohen's d was 0.61, a medium effect based on Cohen's (1992) guidelines.

This is not an unexpected result as it seems reasonable to assume students who have had previous exposure to physics in a high school course, even if they did not do well in that course, would do better on clicker questions than students who have had no previous exposure. This does, however, help establish a higher level of content expertise with the P group than the N group for the purposes of the next two research questions.

Research question 2 analysis. A one-way ANOVA was used to explore the relationship between previous subject experience and SRS correct response rates after peer discussion for each SEP. It was hypothesized discussion pairs who both had previous subject experience (PP group) would more likely answer the clicker questions correctly after peer discussion compared to the NN group where neither partner had previous subject experience. It was also hypothesized the PN group would perform better than the NN group since at least one member of the discussion pair had previous subject experience and the P was more likely to have the correct answer, as previously established.

Because the focus of this research question was dependent upon the interaction of the discussion pairs, partners of students previously eliminated from the study due to dropping or low attendance were also removed from this analysis. This removed six additional students from the analysis.

As noted in Table 4.2 the distributions of the P and N groups were sufficiently normal for the purpose of conducting an ANOVA (i.e., skew<|2.0| and kurtosis<|2.0|). The assumption of homogeneity of variance was satisfied according to Levene's *F* test, $F(2,104) = .785$, $p = .459$.

Table 4.2

Descriptive Statistics Associated with Correct Response Rates after Peer Discussion

Group	N	M	SD	Skew	Kurtosis
PP	24	77.07	7.53	-.091	-.479
PN	46	72.37	10.62	-.528	.495
NN	37	70.58	9.11	-.586	-.360

There was a statistically significant effect of correct response rates after peer discussion at the $p < .05$ level for the three conditions [$F(2,104) = 3.486$, $p = .034$]. The results suggest previous experience with the subject matter may influence the ability to reach a correct answer after peer discussion.

Post hoc comparisons using the Tukey's HSD showed the PP and the NN group differed significantly ($p = .028$) but the PP and the PN group was not statistically significantly different ($p = .126$) nor was there a statistically significant difference in the PN and NN groups ($p = .671$). The results comparing the PP and NN groups supported the hypothesis for this research question but the results comparing the PN and NN groups did not. It seemed reasonable to assume since the P member was more likely to have the right answer prior to the peer discussion, their knowledge would transfer to the N member of the group resulting in a higher percentage correct response rate than the NN group where no such benefit existed.

In comparing the correct response rates before and after peer discussion it is clear all SEP improved with all groups achieving at least a 70% rate after partner discussion. These results are in line with previously established research on peer instruction indicating the largest improvement in percent correct scores occurs when the initial percentage of correct answer is around 50% (Mazur 1997).

Research question 3 analysis. A one-way ANOVA was used to explore the influence of SEP membership on exam performance for the two unit exams administered during this study. It was hypothesized the PP group would have higher exam scores not only because of their previous subject experience but also because they achieved a higher correct response rate on the clicker questions which were representative of questions found on each exam.

As noted in Table 4.3a the distributions of SEP membership were sufficiently normal for the purpose of conducting an ANOVA (i.e., skew<|2.0| and kurtosis<|2.0|).

The assumption of homogeneity of variance was satisfied according to Levene's *F* test, *F*(3,103) = 1.627, *p* = .188.

Table 4.3a

Descriptive Statistics for Exam 4

Group	N	M	SD	Skew	Kurtosis
P of PP	22	79.22	11.83	-.64	-.98
P of PN	25	77.14	13.98	.15	-1.17
N of PN	24	74.41	9.76	.23	-.87
N of NN	36	77.84	11.74	-.59	.36

The results of the ANOVA indicated no statistically significant difference in Exam 4 scores for the four membership conditions, *F*(3,103) = .685, *p* = .563.

Similar results were found for Exam 5. As noted in Table 4.3b the distributions of SEP membership were sufficiently normal for the purpose of conducting an ANOVA (i.e., skew<|2.0| and kurtosis<|2.0|). The assumption of homogeneity of variance was satisfied according to Levene's *F* test, *F*(3,103) = .361, *p* = .782.

Table 4.3b

Descriptive Statistics for Exam 5

Group	N	M	SD	Skew	Kurtosis
P of PP	22	77.00	12.95	-.126	-.696
P of PN	25	78.55	12.06	-.208	-.851
N of PN	24	76.89	14.38	-.487	-.210
N of NN	36	73.57	13.11	-.295	-.559

The results of the ANOVA indicated no statistically significant difference in Exam 5 scores for the four membership conditions, $F(3,103) = .796$, $p = .499$.

These results indicate no lasting influence of previous subject experience or clicker correct response rate. Whatever disadvantage may have existed within the NN peer discussion pair in terms of correct response rates on questions, ultimately did not affect the academic performance for the group on each unit exam.

Research question 4 analysis (MPEX and Partner-Perception Survey). The MPEX survey measures students' cognitive expectations of a physics course as related to their understanding of the process of learning physics and the structure of physics knowledge. The hypothesis was students who had taken high school physics (P groups) would be more in agreement with experts on the MPEX items than students who had not taken high school physics (N group) and that experience or slight "expertise" with the subject could help explain the higher response rates of the PP group over the NN group but also help explain the response rates results of the P of the PN and N of the NN.

For the pre-MPEX scores the hypothesis was students who had taken a high school physics would have a higher "percent agreement with experts" than students who had not taken high school physics. An independent samples *t*-test was used to examine the MPEX scores of "percent agreement with experts" of the P students (M = 31.02%, SD = 12.36%) and N students (M = 27.22%, SD = 13.11%). As noted in Table 4.4a the distributions of the P and N groups were sufficiently normal for the purpose of conducting a *t*-test (i.e., skew<|2.0| and kurtosis<|2.0|). In addition the assumption of homogeneity of variance was tested and satisfied according to Levene's F test, $F(91) = .143$, $p = .706$.

The independent samples *t*-test indicated no statistically significant effect, $t(91) = 1.428$, $p = .157$. These results suggest that while students with previous subject experience are more likely to get to a correct answer on clicker problems and conceptual questions about the content, they have a similar level of understanding or misunderstanding of the process of learning physics and the structure of physics knowledge as those students who never took a high school physics course. This may have implications as to the categorization of students into P and N groups. Content knowledge and process may not be equal.

Table 4.4a

Descriptive Statistics Associated with Pre-Course MPEX Scores

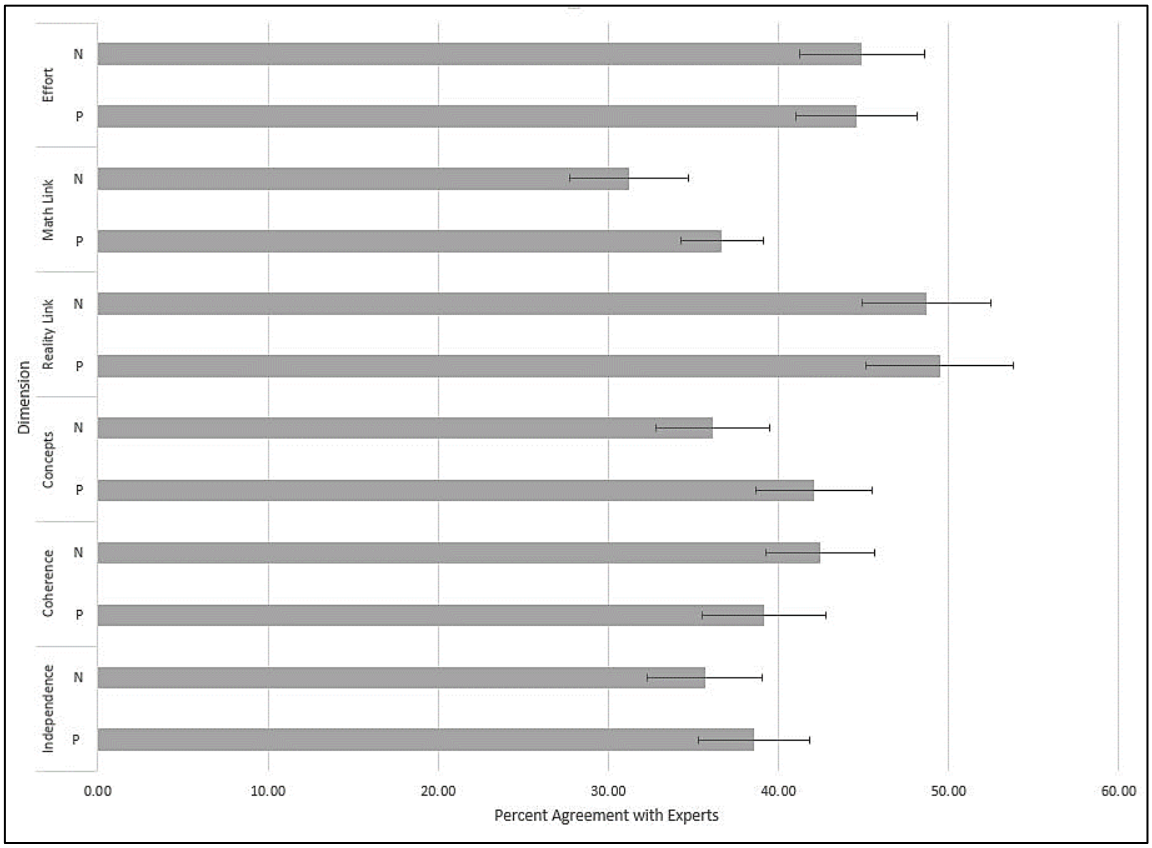
Group	N	M	SD	Skewness	Kurtosis
P	42	31.02	12.36	.303	-.091
N	51	27.22	13.11	.782	1.01

The MPEX is composed of six dimensions and while the overall pre-MPEX scores were not statistically significantly different, it is possible to have an effect on one or more of the individual dimensions. Figure 4a provides a summary of the scores for the P and N groups on each dimension.

As indicated in the chart, there was substantial similarity in results in all six dimensions with the standard error of the mean overlapping in each dimension. No further statistical tests were run on the data. This supports the conclusion from the overall pre-MPEX scores indicating there is no statistically significant difference in “percent agreement with experts" between participants who had high school physics and those who did not.

Figure 4a.

Average Percent Agreement with Experts for Pre-MPEX. Error bars represent standard error of the mean.



The MPEX was administered again at the end of the semester with similar analyses as the pre-MPEX. Additionally, a repeated measures ANOVA was conducted to compare the pre-MPEX and post-MPEX scores. Students who dropped the course or had low attendance, as previously defined, as well as their partners were eliminated from this analysis.

An ANOVA was used to examine the MPEX scores of "percent agreement with experts" of each SEP - P of PP (M = 40.36%, SD = 15.85%), P of PN students (M = 40.24%, SD = 14.65%), N of PN (M = 44.28%, SD = 11.52%) and N of NN (M = 42.16%, SD = 19.87%). As noted in Table 4.4b the distributions of the SEP groups were sufficiently normal for the purpose of conducting an ANOVA (i.e., skew<|2.0| and kurtosis<|2.0|). In addition the assumption of homogeneity of variance was tested and satisfied, Levene's *F* test, $F(3, 78) = 2.834$, $p = .054$.

The results of the ANOVA indicated no statistically significant difference in "percent agreement with experts" for the four membership conditions, [$F(3,78) = .264$, $p = .851$]. These results suggest participants from each of the four SEP had a similar level of understanding of the process of learning physics and the structure of physics knowledge at the end of the semester. It must be noted the MPEX is part of this department's assessment plan and was given at the beginning and end of the semester, meaning the pre-test was outside the scope of this study. Since the last part of the semester involved the peer instruction, however, this researcher felt the immediacy of the experience may be reasonable to influence the post-MPEX results.

Table 4.4b

Descriptive Statistics Associated with Post-Course MPEX Scores

Group	N	M	SD	Skewness	Kurtosis
P of PP	18	40.36	15.85	.06	-1.14
P of PN	22	40.24	14.65	.24	-.50
N of PN	18	44.28	11.52	.05	-.32
N of NN	24	41.71	19.87	.09	-.70

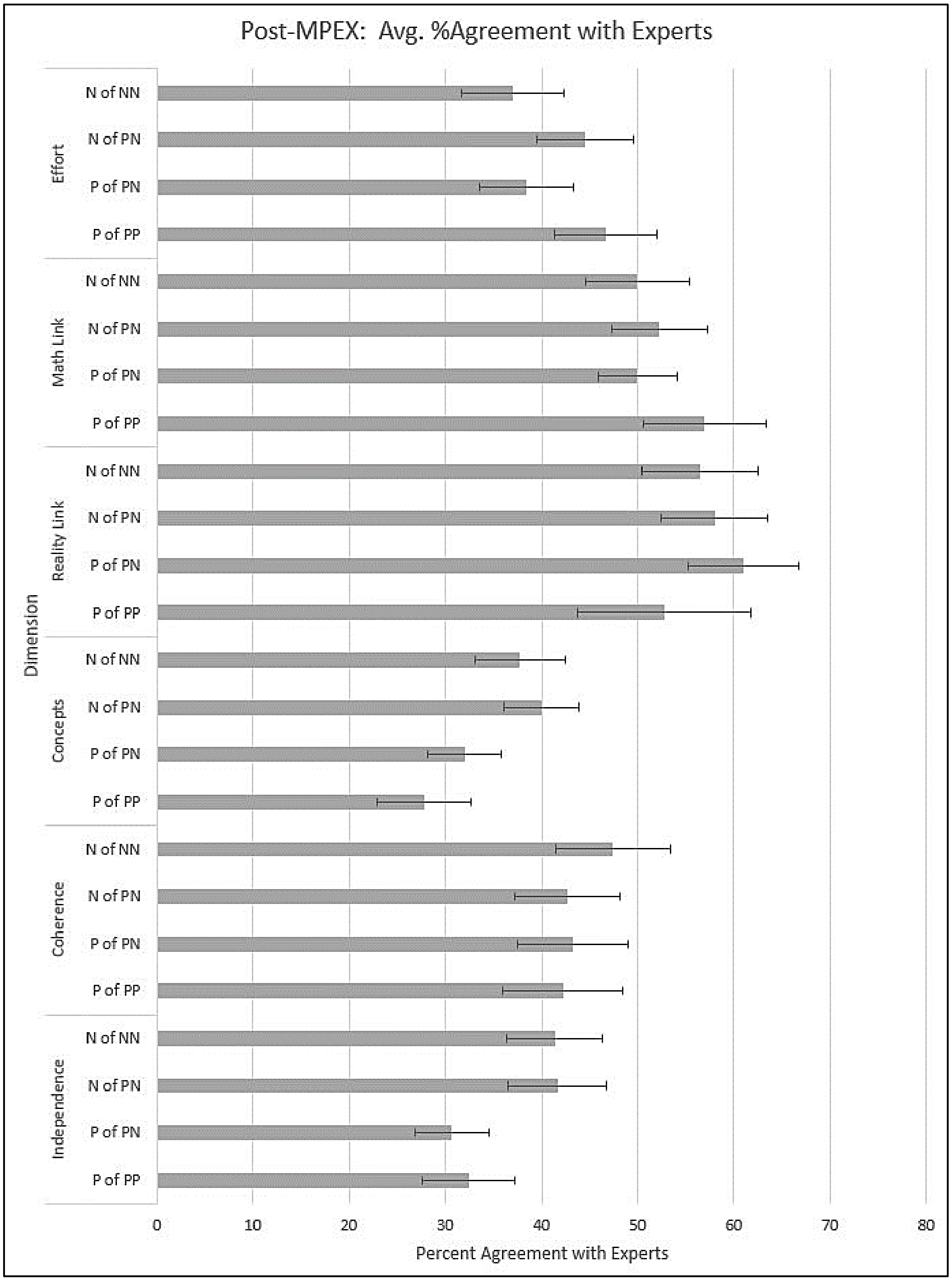
The MPEX is composed of six dimensions and while the overall post-MPEX scores were not statistically significantly different, it is possible to have an effect on one or more of the individual dimensions. Figure 4b provides a summary of the scores for each SEP on each dimension.

Since the error bars represent the standard error of the mean, a visual inspection was used to determine if deeper analysis was needed. The dimensions of Concepts and Independence were further analyzed using a one-way ANOVA. No statistically significant difference was found in the dimension of Independence, $F(3, 81) = 2.03$, $p = .116$ or in the dimension of Concepts, $F(3,81) = .273$, $p = .844$.

A factorial repeated measures ANOVA was used to compare the pre/post MPEX scores for each SEP. The hypothesis was that all SEP members would show an increase in scores over the semester but the N of the PN would benefit from shared experience of a more knowledgeable partner and show greater growth in "agreement with experts" than those in the NN group.

Figure 4b

Average Percent Agreement with Experts for Pre-MPEX. Error bars represent standard error of the mean.



The mean and standard deviation for the pre-MPEX and post-MPEX scores are provided in Table 4.4c. Levene's test for equality of error variances was met for the pre-MPEX scores [$F(3, 75) = .195$, $p = .899$] and was met for the post-MPEX [$F(3, 75) = 2.705$, $p = .051$]. Box's test of equality of variance was met [$F(9, 52931) = 1.089$, $p = .366$]. Since Levene's test was met for both measures, Wilks' Lambda was used to evaluate the interaction of SEP with time. There was no significant interaction of SEP and time, Wilk's Lambda = .986, $F(3, 75) = .357$, $p = .784$, allowing for an analysis of the main effect.

Table 4.4c

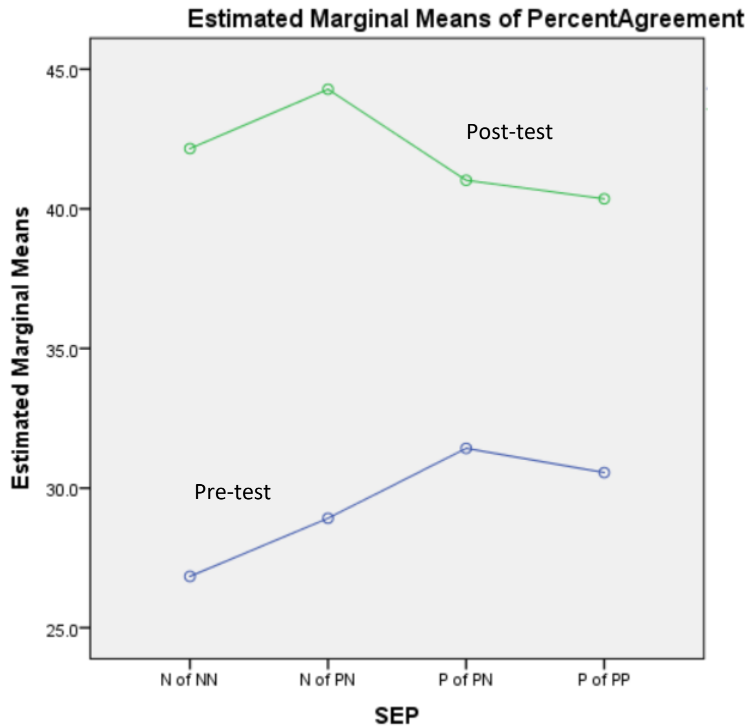
Descriptive Statistics Associated with MPEX Scores

		Pre-Test		Post-Test	
Group	N	M	SD	M	SD
P of PP	18	30.56	13.43	40.36	15.85
P of PN	19	31.42	12.78	41.02	15.08
N of PN	18	28.92	11.95	44.28	11.52
N of NN	24	26.84	14.70	42.16	19.87

The main effect of time was statistically significant, Wilk's Lambda = .782, $F(1, 75) = 20.90$, $p < .000$, indicating there was a significant change in scores from pre-test to post-test (Figure 4c). The partial eta squared was .218 which is considered a large effect according to Cohen (1988). The between-subjects effect was not significant, $F(3, 75) = .245$, $p = .864$, indicating there was not a statistically significant difference between the SEP in MPEX improvement from the pre-test to the post-test.

Figure 4c

Comparison of pre-MPEX and post-MPEX scores.



Participation in the Survey of Physics course did have an effect in improving students' level of understanding of the process of learning physics and the structure of physics knowledge as measured on the MPEX but the nature of the discussion pair did not influence the change in MPEX scores. Although the P of the PP and P of the PN had previous subject-matter experience, they received a similar score to the N of the PN and the N of the NN on the pre-MPEX. This alone may account for the lack of a statistically significant change from pre-test to post-test for the SEP.

In addition to the MPEX participants completed a Peer Instruction Partner Survey (PIPS), provided in Appendix C, a set of six questions in which participants rated their experience with their discussion partner. Each statement in the survey was assessed using a Likert-scale from 1-7 and completed online within QuestionPro. The software allowed for a continuous sliding scale from 1-7. The responses to each question were combined and averaged into one overall PIPS score for each participant.

A one-way ANOVA was used to explore the relationship between PIPS score and SEP. It was hypothesized the N of the PN would score their partner higher than the N of the NN since the P of the PN usually had the correct answer more often on the individual response to clicker questions and that would provide a perceived benefit to the N of the PN in the overall peer instruction experience.

As noted in Table 4.4d the distributions of the P and N groups were sufficiently normal for the purpose of conducting an ANOVA (i.e., skew<|2.0| and kurtosis<|2.0|). The assumption of homogeneity of variance was satisfied according to Levene's F test, $F(3,68) = .236$, $p = .871$.

Table 4.4d

Descriptive Statistics with PIPS scores for each SEP

Group	N	M	SD	Skew	Kurtosis
P of PP	17	5.16	1.43	-1.07	.917
P of PN	20	4.50	1.48	-.265	-1.46
N of PN	15	4.29	1.67	-.003	-.695
N of NN	20	4.71	1.57	-.756	-.085

The results of the ANOVA indicated no statistically significant difference in PIPS scores for the four membership conditions, $F(3,68) = .973$, $p = .411$.

Phase II - Qualitative Analysis

The qualitative phase of the study used focus groups and surveys to explore the nature of the peer instruction experience within each SEP. A separate focus group was held for each SEP, maintaining a commonality of experience for the purposes of discussion. Focus groups varied in size from 5-8 members. The results of the qualitative analysis was used to answer the research questions,

1. What is the nature of previous clicker use or peer instruction experience in each SEP? How might this previous experience have influenced the peer instruction and the quantitative results?
2. How do students in each SEP perceive the value of clickers in learning physics and how might their perceptions explain influence the nature of peer instruction and the quantitative results?
3. How do students in each SEP perceive the value of peer instruction in learning physics and how might their perceptions explain the quantitative results?
4. How do students in each SEP describe the nature of discussion during peer instruction and how might their perceptions explain the quantitative results?

Discussion of Themes

The focus group interviews and survey results provided thick, rich descriptions of participant perceptions and experiences resulting in common and disparate results. Themes developed from the qualitative analysis included: previous experience, perceived value, comfort, and communication. These themes highlighted both external and internal influences on the peer instruction experience.

The research questions shall be discussed in context of the themes developed from the coding process of the data rather than as individual questions.

Previous Experience

The vast majority of respondents had used clickers in a least one class prior to this one and many reported using clickers in two or three prior classes. Most had used the clickers to answer questions either sprinkled throughout lecture, as in this study, or to answer quiz/review questions at the end of a class. This prior experience helps establish a common experience with the technology across each SEP helping to remove the technology itself as a confounding variable in SEP experiences.

Some students expressed a preference for the low-stakes grading method of clicker use in this class over the high-stakes grading they had experienced in others.

> I have used clickers in many Bio and Chem courses. They were used as participation and attendance points as well as others were used as daily points. I find more value in the clickers if they are only used for attendance and participation and cannot actually be detrimental to your grade like they were in (course name).

While the complaint about high stakes grading referenced just the use of the clicker and not peer interaction in previous courses, it does have implications for peer instruction as documented in the literature where it was found if students were given little or no credit for incorrect responses during peer discussion, the peer partner with greater knowledge dominated the conversation (James, 2006). The lack of high stakes grading associated with clicker use and peer instruction in this class was favored by students in each SEP and removed grading as a confounding variable in peer discussions.

Respondents reported little to no previous experience with peer instruction or any type of student-student interaction in conjunction with clicker use. While a few students in each SEP related having conversations with classmates to answer clicker questions in other courses, the vast majority reported individualized use in the context of quiz or review questions.

> Pretty much every class that has over 75 students uses a clicker, but this was the first time I was able to talk to somebody about which answer I got. I really like using clickers because it allows me to assess how well I know the information that I'm being taught and explaining it with a partner helps even more.

Having partner discussion in this course was a new experience for the majority in each SEP and having a dedicated discussion partner was new for everyone.

Perceived Value

With rare exceptions, respondents across each of the four SEP found value in the use of clickers during lecture. The value associated with clickers fell into two categories – engagement and self-assessment/feedback. Many comments centered on how the use of clickers kept them motivated to go to class and helped them stay focused during the class itself.

> I believe clickers keep student involved in the material and attentive. Without clickers I feel attendance goes down along with attentiveness.

Student comments in this study were similar to what has been reported in the literature. Numerous survey results in a variety of classrooms demonstrated students found the use of clickers beneficial in their learning, providing both motivation and engagement (Bazen, & Clark, 2005; Beekes, 2006; Blood & Neel, 2008; Yourstone, Kraye & Albaum, 2008; Salemi, 2009; Stowell & Nelson, 2007; Trees & Jackson 2007;). This positive benefit appeared in a similar number of descriptions across all four SEP.

Additionally, respondents across all four SEP found the interjection of clicker questions throughout a class to be beneficial in helping them assess their learning “on the spot”, providing the desired feedback about their state of understanding of the material.

> I liked these clicker questions because it was a very good way for me to test what I know and if I’m doing the problems correctly.

This method of clicker use was preferred over having all the questions at the end of a class in a quiz or review format.

> I thought it was useful because throughout the lecture we’d be talking about something and then you’d have a question like applied to what we were talking about instead of having questions at the beginning or end which is what I’ve had in a lot of other classes and they use it for like either a review or like “what did you learn today?” kind of thing. I like it more like supplementing what we were talking about.

This perceived value of feedback associated with clicker use has been well-documented in the research literature (Doucent, Vrins, and Harvey, 2009; Schmidt, 2011; Hoekstra & Mollborn, 2012). The benefit of this self-assessment was similarly described across each of the SEP.

These first two themes helped to establish a commonality of experience and perceptions across the SEP. While there was some dissent from the majority in each group, overall participants reported previous experience with the clicker technology and felt the use of clickers in the classroom was valuable for both motivation and feedback. The similarity of these themes across the SEP reduces the chances of their influence in explaining any differences between the SEP.

The agreement among the SEPs diverged on the themes of comfort and communication and on occasion were contradictory especially in the PN group. These divergent themes were related specifically to the peer interaction.

The theme of engagement reappears during the peer interaction but in a different context. While respondents felt the clicker helped them engage with the content and stay engaged in the class (an internal factor) during peer discussion, respondents reported varying levels of engagement between discussion partners, an external factor. This lack of engagement varied in extremity from a partner who was absent quite often to a partner who was simply seemed unwilling to discuss the questions.

> I don't even know how to summarize how my partner and I were 'cause we were pretty relaxed…like we got along pretty well but she didn't seem to care…When we were first assigned partners she had a bad attitude about it (N of PN). She was like, "You can't take me away from my friends, like this is my social hour!" But like when we do our questions I'd be like sitting there doing our individual part of it and like writing out my stuff and figuring it out and she'd just sit there and then we'd go to talk about it with our partners she'd be like "Yeah, I didn't get the answer" and I'd be like, "Well, if you'd one it, maybe you would have!"…so I kind of ended up doing all of it.

Some of the lack of engagement was actually due to the commonality in question answer. If partners had the same answer to a question when they answered individually, often there appeared to be little to no discussion between partners.

> Most of the time we had the same answer and ask, "What you got?" and if we had the same, we don't discuss anything. Just like a couple of times we did get different answers and we discussed how we got different answers so it was misusing a square root or we just didn't convert meters to centimeters or something like that.

If partners did not get the same answer, there was a greater tendency for some level of discussion between partners. Hence, the difficulty of the question was a factor in the complexity of discussion, another external factor.

Comfort and Communication

The themes of comfort and communication emerged primarily from the question where participants were asked to use one word to describe their partner experience and the follow-up questions about the nature of the partner experience. The distribution of the nature of the emotion associated with the one-word response is provided in Table 5. The most common positive response among all four SEP was "helpful or beneficial". Common neutral responses were "ok", "alright" or "average". There was less commonality among the negative response but they varied from "terrible" to "forceful" to "challenging" and "pointless".

Table 5.

Distribution of nature of the emotion associated with the one-word description of the partner relationship during peer instruction for each SEP. The number of respondents includes the focus groups participants and (respondents from the survey).

Group	N	Nature of Emotion Associated with Word		
		positive	negative	neutral
P of PP	5 (16)	71.5%	19.0%	9.5%
P of PN	6 (20)	65.4%	26.9%	7.7%
N of PN	8 (15)	47.8%	39.1%	13.1%
N of NN	6 (19)	44.0%	28.0%	28.0%

A word cloud of each of the group responses is provided in Appendix D. While there is primarily positive emotion associated with the partner experience among the PP and P of PN respondents, the experience was markedly more negative for the N of the PN and the NN groups. The N of the PN participants reported the highest percentage of negative descriptors, and the N of the NN the highest percentage of neutral descriptors.

In the following paragraphs, the themes of comfort and communication are discussed in terms of each SEP partner experience, as related from the focus group interviews and in the survey responses. Since the themes of comfort and communication are sometimes intertwined, this researcher felt it more appropriate to explore these themes in the context of the peer discussion experience for each SEP rather than as individual themes.

The NN Peer Discussion Experience

The majority of negative comments from the NN group related to the assigned seating and not being able to stay with a friend or someone they knew in the class. These comments indicated a lack of comfort with a stranger for peer discussion. The following quote is representative of the flavor of the negative emotions associated with the partner experience for NN group members:

> Again, I don't do well with people in that situation that I don't know. I found myself more distracted texting and talking to the friends I did have in the class after we were split up because I have such bad anxiety when put in situations like partners that I don't know. I won't ask someone I don't know for help so not being able to sit by my friends that could actually help me actually hindered me in this class.

This respondent represents the more extreme end of the spectrum of discomfort with the partner experience but several participants commented on resenting being assigned seating in the class and having a partner they did not know. This was more pronounced in the NN group than any other SEP. Even in situations of positive experience reported by a member in the NN group, the theme of comfort appeared as a confounding factor,

> I had a really positive experience with my partner, I'm pretty quiet and don't like branching out but I was very lucky in having a partner who was positive in helping me get to an answer if I didn't know it and work it out with me.

Both positive and negative experiences appeared to be related to the willingness of partners to communicate and the nature of that communication. If both partners were receptive to the experience and comfortable with sharing information, respondents reported a positive experience.

> I choose helpful because if I did not understand how to do a problem they were always willing to help me understand how to get the right answer and if we both didn't get it we would work together to try and come up the best answer.

For some participants this discomfort with an unknown classmate was a motivating factor pushing the student to do their best so they didn't look "stupid" or "foolish" in front of someone they didn't know.

> I think having an assigned partner really kinda made me do my best on each question 'cause I'm like I don't want them to think an idiot and so I like did my best and if I absolutely could not get it my partner would typically help me or like if either of us didn't get it we'd like try to work together. So having that partner really just kinda reinforced, uhm, my work that I was doing and then I was able to help her too which was nice.

The majority of respondents perceived their partner to have fairly equal physics content knowledge to their own. Since participants were not told how the partner groupings were arranged, this perceived equality came from the peer interaction itself. In situations where partners did not have the same answer, some respondents described a lack of discussion due to lack of expertise.

> Yeah, if we'd agree, it'd be, "B". "Ok, cool." Then if we didn't…neither of us knew what we were doing, "Well, yours makes more sense than mine so we'll go with that. Ok, let's hope that's right."

There were, however, situations where the disagreement in answers led to a productive discussion, as in other SEP. Situations where both members of the group had the same answer resulted in little to no discussion but for members of this group, respondents were more likely to describe poor communication or a sense of helplessness in trying to get to the right answer.

Overall, the NN participants reported a more positive than negative experience. Negative experiences appeared to be mostly related to forced seating and partner assignment which resulted in both a lack of comfort in the peer discussion experience and a reduced level of communication.

The N of the PN Peer Discussion Experience

Although some of the N of the PN members expressed a similar concern about having assigned seating and an assigned partner in the class, this was less of a topic than in the NN group. Dissatisfaction for the N of the PN group was usually more related to the level and type of communication with their partner. Some students complained of their partner being absent too often. While a negative emotion in itself, it does demonstrates the desire to have a partner present to help with the clicker questions.

A few negative experiences reported by the N of the PN members arose from a perceived inequity in content knowledge.

> My partner thought that she literally knew how to do everything. Even if I told her that I got a different answer than her she was like well yours is wrong and then mine would be right and hers would be wrong. She would only talk to me to brag about something awesome she did or whine about stuff. I did not enjoy her what

> so ever, and found myself turning to my other side and conversing with the person on my other side instead.

This same inequity, however, also served as a positive experience for some members of the same group.

> I was fortunate to have a partner that had a much better understanding of the formulas and was willing to show me how he set it up and found the answer.

The reported positive experiences appeared to stem from the ability or willingness of partners to communicate.

The N members perceived their P partner as having either equal or less knowledge about the subject matter. Since participants were not told what criteria used to establish partners, this perceived equity or inequity was a result of the peer interaction.

> …once we were assigned partners, I was placed next to a gal that had no idea what she was talking about when it came to physics, so I never discussed anything with here because we were not on the same page at all.

Although the results of the quantitative analysis showed P students were more likely to have the correct answer than N students, it may be the case that in some partnerships the N learned the material more thoroughly than their peer counterpart or it may be the perceived expertise of the P partner was related to how willing or able the partner was in communicating their knowledge and less related to the actual content knowledge.

Overall, the N of the PN participants reported a more positive than negative experience during peer discussion but the negative experiences were more pronounced than the NN group. Negative attitudes and perceptions appeared to result from both the forced seating and partner assignment and problems with a partner's willingness or ability to communicate during the discussion.

The P of the PN Peer Discussion Experience

The P of the PN participants expressed a more positive experience related to the peer interaction than their counterparts in the same group. Many expressed a benefit of discussing the question with their partner, seeing the interaction as a means of seeing a "different point of view" or providing "someone to bounce ideas off of" or that "it was good to have a resource to explain how to do it".

> Sometimes you just need to talk a question out with someone else. It helps to discuss a tricky problem with someone else and see if they caught something you may have overlooked.

The primary benefits of the peer interaction from the perspective of the P of the PN appeared to relate more to "tricky problems" or "finding small mistakes" rather than relying on a partner to show them how to do an entire problem or explain the answer to a question they completely misunderstood. Descriptions indicated the P of the PN was more comfortable with the nature and content of the questions than the N of the PN.

A large number of respondents in this SEP described their partners as having about equal content knowledge. Considering they had not been told the criteria used to form discussion pairs, this perception must have come from the interaction. This is interesting since members of this group (P of PN) demonstrated a higher level of content knowledge by getting the clickers questions initially correct more often than their counterparts as reported in the quantitative findings. This is somewhat contradictory to the N of the PN who, more often than not, reported their partners as having equal or less knowledge than they did. The perceived content knowledge may have more to do with the level of communication between partners and less to do with actual knowledge. The P of the PN participants who knew more about a question were not necessarily able or willing to communicate that knowledge to their partner. The MPEX results from the quantitative study indicated no difference between the SEP on the understanding of the process and structure of physics, attributes helpful in mentoring or teaching the subject.

As reported in other groups, the P of the PN disliked and were frustrated when their partner was absent, indicating a perceived value in the partner discussion. The members of this group also reported deeper levels of discussion occurred when the partner did not initially have the same answer.

The P of the PP Peer Discussion Experience

This group had the highest number of positive descriptors about the nature of the peer interaction. Very few students expressed a negative attitude about the assigned seating component and the majority of students described a higher level of communication and a greater openness than any other SEP.

> I don't know how to describe it but I guess the person I discussed with…there was never a...I don't know…looking down on anything 'cause I guess we each got 'em wrong a fair amount of time but it was never a…I don't know…it was always acceptable or ok to do that. I mean we each knew we did it about the same amount.

Members of this group tended to describe their partners as "good listeners", "able to communicate", "observant", and "helpful". There were clearly more examples of cooperative assistance and teaching between the PP members whether it was the result of not getting the same answer leading members to "talk about it and try to figure out a solution" or how the discussion "helped generate more ideas which generally led to more correct answers."

As with other SEP members, discussions were limited when both members had the same answer prior to discussion.

> I would say collaborative because we would like check with each other like what we got and if it would be the same that would be great or it'd be like ok, one of us did something wrong…what did we do? What can we do to fix it?

Lengthier, more involved and meaningful discussions occurred when there was a conflict in answers or approach.

Most members of this groups perceived their partners as either equal in content knowledge or one partner was perceived as having greater content knowledge. In situations of a described inequity, there was generally a positive attitude and experience associated with it.

> I was paired with a person who obviously understood it better than I did so they were teaching me. I was kinda the opposite side of him and so I would always guess. I would answer mine and the he would ask, "What did you put?". He would never tell me the answer and then he'd have me explain why I'd put it and then he'd either tell me, "Yea, that what I put" or he would say, "Hmmm, you're thinking of like that other like series or something like that. This is probably what it should...probably is" and he was mostly right.

The described experience for the PP members was markedly different than the other SEP. Both communication and comfort were higher in this group. Whether that stems from a higher level of comfort with the subject matter for both partners or the perception of equity between partners is unclear.

Chapter 5

Discussion

This study utilized an explanatory sequential mixed methods design, a two-phased design in which qualitative data helped explain or build upon initial quantitative results. For this study the mixed methods approach utilized the "Follow-up Explanations" sub-model of the explanatory design (Creswell & Plano Clark, 2011) with the following mixed-methods research question as the focus:

> What is the difference in experiences in a college Survey of Physics course between peer instruction discussion partners who have no previous subject experience compared to student discussion pairs in which only one or both students have previous subject experience?

The quantitative data was the primary focus of this study with the qualitative data in a supportive, explanatory role as indicated in Figure 5.1.

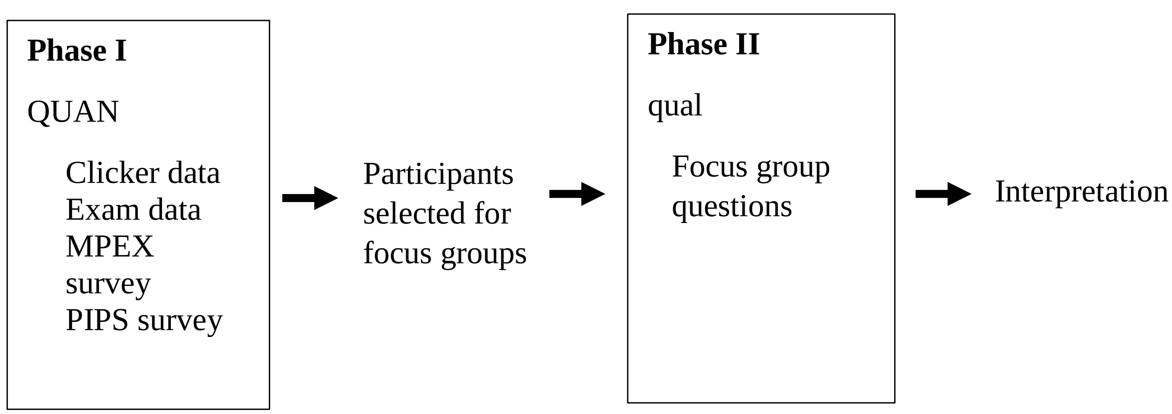


Figure 5.1. Summary of follow-up explanatory model of mixed methods research used in this study.

Development of Qualitative Questions from Quantitative Results

The quantitative results indicated the P participants were more likely to answer clicker questions correctly when first asked individually, helping validate the hypothesis of an initial benefit in content expertise for students with previous subject experience. This apparent benefit of previous subject experience appeared to make a difference after peer discussion as well since the PP groups were statistically more likely than the NN groups to reach the correct answer after discussion. Members of both groups were in the same Survey of Physics class exposed to the same material with the same instructor suggesting previous subject experience may be a contributing factor in the correct response rates after peer discussion. Since peer discussion involved communication between assigned partners, there was a need to explore the nature of this interaction to determine if previous subject experience was the sole factor in the difference in correct response rates. This was explored in the qualitative part of the study.

The correct response rates of the PN group was not statistically different from the PP or NN groups after peer discussion. This was a contrary to this researcher's hypothesis that having someone with previous subject experience to discuss content questions should prove beneficial to someone without previous subject experience. From a qualitative perspective it becomes natural to wonder why the benefit did not express itself. Was there something about the nature of the PN interaction preventing the N of the PN from doing better than the N of the NN?

The documented "expertise" of the P participants, other than just having taken a high school physics course, is limited to their responses on conceptual and mathematical clicker questions. The results of the MPEX survey, which measured a participant's level of understanding of the process of learning physics and the structure of physics knowledge, indicated no statistically significant difference between the four SEP. This means students who had a full year of high school physics were not statistically different on this measure than students who had no previous experience with the subject. While beyond the scope of this study, it would be of interest to explore this result further as it may be related to the teacher preparation of high school physics teachers who are primarily, especially in rural states such as the one in this study, trained biology teachers rather than teachers trained in the content discipline while in college.

The MPEX results may provide insight into the post-discussion clicker results. Content “expertise” does not necessarily equate with process and structural physics knowledge which may be characteristics necessary to serve in a mentoring-type role such as that hypothesized of the P of the PN group. The qualitative study was used to explore the peer discussion interactions in an attempt to reveal the nature of the communication between members of each SEP which may support this conclusion.

Previous subject experience did not lead to an advantage on the unit exams as the average scores for each of the two unit exams was not statistically significant among the four SEP. This may imply exposure to the topics through lecture and lab experiences in this class was enough to achieve similar success by the time of the exams. This result was not explored any further in the qualitative aspect of the study.

The PIPS results were not statistically significant between the SEP with all groups providing a PIPS score above average. The qualitative analysis allowed elaboration on the peer discussion experience with some contradictory results.

Using Qualitative Results to Explain Quantitative Results

While there were several quantitative research questions, not all of them required a qualitative counterpart for further analysis. For example, the statistical result indicating the P participants were statistically more likely to get the first clicker response correct was used to establish an “expertise” level for these members and was not considered for further qualitative analysis. This was the case with several other quantitative questions.

The purpose of the qualitative study was to draw out information to help explain the unexpected or contrary quantitative results. These qualitative questions focused primarily on the partner interaction. The relationship between the qualitative and quantitative results are presented in Table 5.1 in a “follow-up results joint display” (Creswell, 2015).

Table 5.1

Connection of Quantitative and Qualitative Data to Explain which SEP Factors Influence Peer Instruction Results

Quantitative Results	Qualitative Follow-Up Interview Results Relating to Quantitative Results	How Qualitative Findings Helped to Explain the Quantitative Results
The results of the MPEX were not statistically significantly different across the SEP	• P of the PP usually perceived their partners as having equal or greater content knowledge • P of PN usually perceived their partners as having equal content knowledge • N of PN usually perceived their partners as having equal or lesser content knowledge • N of NN usually perceived their partners as having equal content knowledge	Although P participants expressed greater content expertise (as measured on first-response clickers questions), partners did not always perceive this greater knowledge. This may be related to the P member’s ability and willingness to communicate their knowledge, their level of understanding of the material presented in a given unit, or the interpretation of the explanations by other partner

The PP group achieved statistically significantly higher correct-response rates on clicker questions after peer discussion than the NN group.	The NN group members expressed a higher level of frustration with the peer discussion experience related to the comfort in discussing questions with a stranger and the ability or willingness of the partner to communicate	The nature of the communication and the level of comfort in conversing with a stranger was notably different between each SEP. In some cases this negative attitude or emotion nullified any academic benefit of peer discussion.
The PN group correct-response rate on clicker questions after peer discussion was not statistically significantly different than the PP or NN group.	The P of the PN group reported a more positive experience than their N counterpart. While the P was sometimes seen as having greater or lesser knowledge than the N partner, more often the N reported the P as having equal content knowledge.	Communication between the P and the N was more strained than that reported in the PP group. Often this was related to perceived or real ability of one partner to communicate with the other.
Participants rated their peer discussion partners similarly on the PIPS Survey.	The N of the NN and the N of the PN had greater negative emotion associated with their one-word descriptors of their partner experience than the PP or P of the PN participants	Providing one overall PIPS score masked an underlying notable difference in experiences between the SEP, especially the P versus N group members

Although P participants were more likely to correctly answer clicker questions when first asked, it is unclear if this higher level of initial content expertise was truly relevant or advantageous for the N participant within the PN pair. The quantitative results indicated no statistically significant difference of the PN group in getting the correct answer on clicker questions after discussion as compared to the PP or NN groups whereas the PP groups did achieve a higher correct rate after discussion than the NN group. The qualitative data provides insight into why the PN pair may not have achieved the same level of success as the PP. Within the PN pair, N participants typically perceived their partners as having equal or lesser content knowledge. This perception is likely related to the ability or willingness of the P partner to communicate, something that clearly did not happen as well within PN pairs as PP pairs. Having the content knowledge and effectively expressing the content knowledge are not necessarily the same thing and the response rates measured in the quantitative part of the study only measured content knowledge.

The MPEX scores provided some insight into participants' ability to understand the underlying structure of physics knowledge. Perhaps these characteristics are necessary to explain content to someone else. Since there was no statistically significant difference in MPEX scores between the SEP, the ability to explain physics to another person may not have been different between the PN partners, especially the subtleties related to several of the clicker questions. The advantage of the content knowledge may have been effectively negated by the inability or unwillingness to share it.

Within the qualitative study the nature of the interactions was more clearly described between the PN members. While a few of the N partners described their P counterpart has having greater content knowledge and providing instructive guidance on clicker questions, it was more often the case the N partner described their partner as having equal or lesser content knowledge and not being helpful during the peer interaction either due to an unwillingness to discuss questions or a perceived inability to be helpful.

The P of the PN participant was more likely to express a positive experience in the interaction and so perhaps their opportunity to help made them feel better whether or not it was effective. The N did not share this same feeling, however. The effort of the P was usually perceived as either overt condescension or not helpful by their N counterpart, with some exceptions. Perhaps the explanation provided by the P member was not coherent from the perspective of the N participant, thereby leading them to conclude their partner was less knowledgeable even though the P partner was correctly communicating content knowledge. Whatever the case, the communication level between the P and the N pair was not as productive as that described by PP members. There was less complaint about the forced seating and assigned partners from the N of the PN than the NN indicating some positive interaction and experience associated with the N and P pairing.

For the majority of NN participants there was a strong resentment about the forced seating arrangement and assigned partner for this study. Since this was expressed more strongly in the NN pairing than from the N of the PN participants, there must have been some perceived benefit for the N of the PN which did not express itself in the NN grouping. There was no statistically significant difference in PIPS scores for the four SEP but combining rankings on all six statements into one score was clearly masking hidden problems which only revealed themselves during the qualitative part of the study. NN members were much more likely to express a discomfort with having a peer discussion partner they didn't know. This discomfort appears to have led to an unwillingness or an inability to effectively communicate with the partner during peer discussion.

Implications

This study has provided insight into the nature of peer discussions between students with varying levels of previous subject experience. The major contributions of this study were to determine if the level of previous subject experience had short-term (response-rates on clicker questions) and long-term (exam scores, MPEX scores) effects. This study has reinforced some of the existing literature showing the value of the use of peer instruction in answering clicker questions during class since the majority of students achieved the correct answer to these clicker questions after peer instruction. This study adds to the current literature by expanding the knowledge about the interactions and attitudes between students during peer instruction, especially as it relates to forced partnerships and the influence of previous subject experience.

The results of this study are intended for teachers using peer instruction in the physics classroom. While the value of clickers and peer instruction has been well-established in the literature in a general sense, there are subtleties in implementation that need to be considered. About half of students entering a college-level, non-majors physics will likely not have taken high school physics. These students are normally apprehensive about the subject-matter and may try to increase their comfort level in the course by taking it with friends. If they are then removed from their friends and forced to sit with a stranger for the purposes of peer instruction, that comfort level has been removed. This lack of comfort may dissipate as the semester progresses especially if peer instruction partners are willing to openly communicate with each other, but for many the lack of comfort remains. Any potential benefit of placing a student who had high school physics with a student who did not appears to be negated by the social-emotional aspect of comfort associated with having a friend as a partner. This is not as pronounced with students who had previous subject experience perhaps because their comfort level in the course starts on two levels – some comfort with the subject matter because they had the course in high school and the added comfort of potentially sitting next to a friend in the class. When these students (P members) are removed from their friends and assigned a stranger as a partner, they still maintain at least a comfort level with the subject matter and this appears to lessen the strain of the experience.

Factors influencing the peer instruction experience were both internal and external and it is the external factors which can be most directly influenced by the teacher. Clicker questions with obvious answers, where both partners are likely to get the same answer, generate little to no discussion between partners. Having deep, rich, complex questions are of greatest value for peer discussion. Assigning partners, another external factor, may negatively influence students to the extent of negating any other content expertise benefit to the peer discussion.

Limitations

This study is limited by nature of the sampling. Participants in this study were drawn from a sample of convenience, this researcher's own Survey of Physics course. While this allowed for control of methodology of clicker implementation and peer instruction, it limited sample size and generalizability to a broader population. Conducting this study across multiple institutions and multiple regions of the country would increase the robustness of the study.

It was difficult to get students to participate in the focus groups. After initial invitations were emailed to students, a fair number declined to participate either due to scheduling conflicts or a lack of desire. As a result, for both the PP and members of the PN groups, every member of the group was eventually sent an invitation just to ensure enough would attend the focus group sessions. This was not true of the NN group where enough students who were initially invited agreed to participate. As a result, the qualitative sample was not truly purposefully sampled for the PP, P of PN, and N of PN focus groups but rather ended up being a more self-selected group of individuals. These individuals may not be truly representative of the group as a whole.

Member-checking of the themes developed in the qualitative study did not occur to the extent desired by this researcher. No members of the PP group responded to the request to review a summary of the focus group session and survey results. Only two members of the NN group responded and only one participant from each of the other two groups provided feedback. All responses provide were quite short.

The MPEX survey was given at the beginning and end of the semester as dictated by the department assessment plan, while this study occurred only during the last third of the semester. While the influence of peer instruction was in proximity to the post-test, it may be more valuable to have the pre-MPEX completed closer to the beginning of the study thereby bracketing the peer discussion experience under study. This would imply the study occur at the beginning of the semester rather than at the end since the pre-MPEX should measure initial perceptions before the class starts.

Recommendations for Future Research

1. This study occurred at the end of the semester covering a unit on light and sound and another unit on electricity and magnetism. It is unclear if the classification of students into P and N groups based solely on whether they had a high school physics class is sufficient to define a level of “expertise”. Consideration needs to be given to the fact that the topics of electricity and magnetism are often short-changed or skipped in high school physics courses. Students could be given a pre-test prior to the start of the study and the results of that pre-test used to define the P and N groups.
2. Does the social-emotional aspect of “sitting with a friend” change the results of this study? In other words, if students are allowed to sit with whomever they

want there would still probably be a useable distribution of SEP pairs. This would remove the “lack of comfort with a stranger” component for many students. Does that change the results of the study?

3. What characteristics are required for successful peer discussion or peer tutoring? Does the MPEX measure any of these characteristics as it relates to physics peer tutoring which may be necessary for successful peer discussion? Should some other measure be used for grouping – one that focuses on characteristics necessary for peer tutoring?
4. How does the content expertise (training) of a high school physics teacher influence MPEX scores? Do students of teachers who are trained in the discipline score higher than students of teachers who were trained primarily as biology teachers?

Appendix A

IRB Approval Letters

Nebraska University of Lincoln

November 29, 2016

Judy Vondruska
Teaching, Learning and Teacher Education

Allen Steckelberg
Teaching, Learning and Teacher Education
59 HENZ, UNL, 68588-0355

IRB Number: 20161116648 EX
Project ID: 16648
Project Title: Exploring the nature of interactions during peer instruction for different subject-experience pairs

Dear Judy:

This letter is to officially notify you of the certification of exemption of your project. Your proposal is in compliance with this institution's Federal Wide Assurance 00002258 and the DHHS Regulations for the Protection of Human Subjects (45 CFR 46) and has been classified as exempt.

You are authorized to implement this study as of the Date of Final Exemption: 11/29/2016.

o Review conducted using Exempt category 1 and 2 at 45 CFR 46.101
o Funding: NA

Please use all documents approved through South Dakota State University in your research.

We wish to remind you that the principal investigator is responsible for reporting to this Board any of the following events within 48 hours of the event:
* Any serious event (including on-site and off-site adverse events, injuries, side effects, deaths, or other problems) which in the opinion of the local investigator was unanticipated, involved risk to subjects or others, and was possibly related to the research procedures;
* Any serious accidental or unintentional change to the IRB-approved protocol that involves risk or has the potential to recur;
* Any publication in the literature, safety monitoring report, interim result or other finding that indicates an unexpected change to the risk/benefit ratio of the research;
* Any breach in confidentiality or compromise in data privacy related to the subject or others; or
* Any complaint of a subject that indicates an unanticipated risk or that cannot be resolved by the research staff.

This project should be conducted in full accordance with all applicable sections of the IRB Guidelines and you should notify the IRB immediately of any proposed changes that may affect the exempt status of your research project. You should report any unanticipated problems involving risks to the participants or others to the Board.

If you have any questions, please contact the IRB office at 402-472-6965.

Sincerely,

[redacted]

Becky R. Freeman, CIP
for the IRB

Office of Research/Human Subjects Committee
SAD Room 124
Box 2201 SDSU
Brookings, SD 57007

To:	Judy Vondruska, Department of Physics
Date:	September 15, 2016
Project Title:	Exploring the nature of interactions during peer instruction for different subject-experience pairs
Approval #:	IRB-1609006-EXM

Thank you for taking such care in completion of the request and research protocol. This project is approved as exempt human subjects' research. The basis for your exempt status from 45 CFR 46.101 (b) is:

(1) Research conducted in established or commonly accepted educational settings, involving normal educational practices, such as (i) research on regular and special education instructional strategies, or (ii) research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods; and
(2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:
(i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.

If there are any unanticipated problems involving risks to subjects or others, or changes in the procedures during the study, contact the SDSU Research Compliance Coordinator. Protocol changes must be approved by SDSU prior to implementation. At the end of the project please inform the committee that your project is complete.

If I can be of any further assistance, don't hesitate to let me know.

Sincerely,
[redacted]

Norman O. Braaten
SDSU Research Compliance Coordinator

Appendix B

Qualitative Interview Documents

Focus Group Script

WELCOME

Thank you for taking time to participate in this focus group. Your willingness to share your opinions is appreciated.

PURPOSE OF THE FOCUS GROUP

The reason for this focus group is to hear your opinions about the use of clickers in Physics 101 and the experience of having a discussion partner to answer some of the questions. Please be open and honest in your responses.

GROUND RULES

1. You do the talking.
    a. It will be most valuable if everyone participates.
    b. I may call on you if I haven't heard from you for a while.
2. There is no right or wrong answer.
    a. Everyone's experience and opinions are important.
    b. Speak up whether you agree or disagree.
    c. A wide range of opinions is desired.
3. What is said in this room stays here.
    a. It is important everyone feel comfortable sharing their opinions.
4. This focus group session will be recorded.
    a. It is important to capture everything you say.
    b. When the recording is transcribed, no one will be identified by name. You will remain anonymous.

QUESTIONS

1. Prior to Physics 101, what experience have you had with clickers in a course? Please identify the course, the way the clickers were used, and your sense of the value of their use in the course.
2. Let's initially separate the use of clickers from the discussion of questions with partners. In other words, let's just focus on the situation where a question was posed and you were asked to respond individually, without talking with your partner. Describe what you did when a clicker question was asked. Was your behavior different with different types of questions?
3. Now let's consider the times you were asked to discuss your answer with your partner. Describe the nature of your interaction with your partner; the nature of the conversation.

4. Some questions asked during the semester had you work with a partner right away while others required you to answer on your own first and then discuss it with a partner. What are your feelings about each of those two methods?
5. What changes would you suggest for the use of clickers in Physics 101 in the future?
6. Is there anything else you would like to add to our conversation today?

Focus Group Invitation (Email)

Dear _________,

You have been randomly selected to participate in a focus group to discuss the use of clickers in Physics 101 and the experience of having a discussion partner to answer some of the questions. The session is scheduled for ________________ and will last approximately one hour. Your participation in this focus group is completely voluntary.

As part of this focus group, you will be with 8-12 other students from class. Although the session will be recorded, your responses will be anonymous. Pizza and pop will be provided during the session and a $10 Wal-Mart gift card will be given to you at the end of the session. If desired, you can also enter your name into a drawing for an additional $50 gift card which will be given to one participant.

Please accept or decline this invitation through email by 8 pm tonight. If you accept, I will send you specifics about the meeting room and other additional information about the focus group.

Thank you for your consideration. I look forward to hearing from you.

Sincerely,

Judy

Focus Group Confirmation Email (Email)

Dear _________,

Thank you for your willingness to participate in the clicker focus group. As described in the original email, I would like to hear your opinions about the use of clickers in Physics 101 and the experience of having a discussion partner to answer some of the questions. As part of this focus group, you will be with 8-12 other students from class. Although the session will be recorded, your responses will be anonymous. Pizza and pop will be provided during the session and a $10 Wal-Mart gift card will be given to you at the end of the session. If desired, you can also enter your name into a drawing for an additional $50 gift card which will be given to one participant. The date, time, and location for the focus group is listed below.

Date
Time

Location **Daktronics Hall – Room 269**

I have attached a consent form you will need to fill out in advance. Please print it off, sign it and bring it with you to the session. If you happen to forget it, I will have extra copies available at the focus group session.

If you have any questions regarding the focus group, please do not hesitate to email me or call, 688-5859. I look forward to seeing you at the focus group session.

Sincerely,

Judy

Consent to Participate (sent via Email or filled out before session begins)

Consent to Participate in Focus Group

You have been asked to participate in a focus group for Physics 101. The purpose of the focus group is to elicit your opinions about the use of clickers in Physics 101 and the experience of having a discussion partner to answer some of the questions. The information learned in the focus group will be used to continue/modify the use of clickers in future classes and is being used a part of a research study for publication on the topic.

You can choose whether or not to participate in the focus group and stop at any time. There is no known risk to your participation. Although the focus group will be tape recorded, your responses will remain anonymous and no names will be mentioned in the research paper.

Your choice to participate or not participate in this focus group will have no impact on your course grade. What you say in the focus group will have no impact on your course grade.

There are no right or wrong answers to the focus group questions. The goal is to hear many different viewpoints and to hear from everyone in the group. Please be honest in your responses even when your responses may not be in agreement with the rest of the group. Please be respectful of everyone in the group by allowing only one individual to speak at a time and by agreeing to keep responses made by all participants confidential.

I understand this information and agree to participate fully under the conditions stated above:

Signed:_____________________________________ Date: ______________

Appendix C

Peer Instruction Partner Survey

Directions: Please rate each of the following statements based on your experience with using clickers in lecture to answer questions and having an assigned partner to discuss the clicker questions. While your name is entered on the survey for initial filing purposes, your name will be removed when the survey is analyzed. Your responses to the survey will remain confidential.

Name:

Statement	Rating						
If found it helpful to work with my partner on clicker questions in class.	1	2	3	4	5	6	7
I feel my partner communicated well with me during discussions.	1	2	3	4	5	6	7
I felt more confident in answering clicker questions when I could discuss it with my partner.	1	2	3	4	5	6	7
My discussion partner positively influenced my attitude about learning physics.	1	2	3	4	5	6	7
Being able to discuss questions with my partner helped me to do better on exams.	1	2	3	4	5	6	7
If I could, I would like to continue to work with the same partner in the next unit.	1	2	3	4	5	6	7

Appendix D

Word Clouds* for One-Word Descriptions of Partner Discussion Experience

P or PP Participants (n = 21)

P of PN Participants (n = 26)

NN Participants (n = 23)

N or PN Participants (n = 25)

*Created at https://www.wordclouds.com

Words used to describe the peer interaction experience and the emotions assigned:

PP	**P of PN**
teaching + helpful + helpful + collaborative + challenging - helpful + good + good + helpful + minimal - challenging - helpful + average n good + interesting + OK n pointless - excellent + useful + productive + beneficial +	professional + awkward - professiona+ forced - beneficial + inconsistent - interesting + relaxed + unexpected + good + helpful + beneficial + educated + interesting + helpful + dumb - helpful + gone - awesome + good + valuable + alright n alright n useless - pointless - informative +
NN	**N of PN**
helpful + challenging - forceful - agreement n reinforcement + obligation - ok - good + average n average n decent n helpful + positive + noncommunicative - helpful + short - eh - okay n horrible n unnecessary n beneficial + beneficial + useful + helpful + beneficial +	short - relaxed + distant - educational + relaxed + informative + short - boring - helpful + careless - average n helpful + nonexistent - indifferent n fine + great + helpful + inadequate - great + snob - OK n helpful + terrible -

REFERENCES

Anderson, L. S., Healy, A. F., Kole, J. A., & Bourne, L. E. (2011). Conserving time in the classroom: The clicker technique. *Quarterly Journal of Experimental Psychology*, *64*(8), 1457-1462.

American Institute of Physics. (2014). [Graph illustration of Physics-Taking in U.S. High School from 1948-2013]. *Physics Trends*. Retrieved from https://www.aip.org/statistics/physics-trends/physics-taking-us-high-schools-1948-2013

Anthis, K. (2011). Is it the clicker, or is it the question? Untangling the effects of student response system use. *Teaching of Psychology, 38*(3), 189-193.

Bazen, E. F., & Clark, C. D. (2005). Promoting interactive learning with an electronic student response system. *NACTA Journal, 49*(3), 11-16.

Beatty, I. D., Gerace, W. J., Leonard, W. J., & Dufresne, R. J. (2006). Designing effective questions for classroom response system teaching. *American Journal of Physics, 74*(1), 31-39.

Beekes, W. (2006). The “Millionaire” method for encouraging participation. *Active Learning in Higher Education, 7*, 25-36.

Beckert, T. E., Fauth, E., & Olsen, K. (2009). Clicker satisfaction for students in human development: Differences for class type, prior exposure, and student talkativity. *North American Journal of Psychology, 11*(3), 599-611.

Berg, B.L. (1998). *Qualitative research methods for the social sciences.* Needham Height, MA: Pearson Education Company.

Berger, C.R. & Calabrese, R.J. (1975). Some explorations in initial interactions and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Theory, 1,* 99-112.

Bibles, B. D. (2011). Use of classroom response systems for formative assessment in natural resource courses. *Journal of Forestry, 109*(7), 417-420.

Blasco-Arcas, L., Buil, I., Hernández-Ortega, B., & Sese, F. J. (2013). Using clickers in class. The role of interactivity, active collaborative learning and engagement in learning performance. *Computers & Education, 62,* 102-110.

Blood, E. (2012). Student response systems in the college classroom: An investigation of short-term, intermediate, and long-term recall of facts. *Journal of Technology and Teacher Education, 20*(1), 5.

Blood, E., & Neel, R. (2008). Using student response systems in lecture-based instruction: Does it change student engagement and learning? *Journal of Technology and Teacher Education, 16*(3), 375-383.

Brewe, E. (2008). Modeling theory applied: Modeling Instruction in introductory physics. *American Journal of Physics, 76*(12), 1155-1160.

Bunce, D. M., Flens, E. A., & Neiles, K. Y. (2010). How long can students pay attention in class? A study of student attention decline using clickers. *Journal of Chemical Education, 87*(12), 1438-1443.

Bunce, D. M., Vandenplas, J. R., & Havanki, K. L. (2006). Comparing the effectiveness on student achievement of a student response system versus online WebCT quizzes. *Journal of Chemical Education, 83*(3), 488-493.

Cain, J., Black, E. P., & Rohr, J. (2009). An audience response system strategy to improve student motivation, attention, and feedback. *American Journal of Pharmaceutical Education*, *73*(2), 1-21.

Carnaghan, C., Edmonds, T. P., Lechner, T. A., & Olds, P. R. (2011). Using student response systems in the accounting classroom: Strengths, strategies and limitations. *Journal of Accounting Education*, *29*(4), 265-283.

Carnaghan, C., & Webb, A. (2007). Investigating the effects of group response systems on student satisfaction, learning, and engagement in accounting education. *Issues in Accounting Education*, *22*(3), 391-409.

Chickering, A. W., and Z. F. Gamson. 1987. Seven principles for good practice in undergraduate education. *AAHE Bulletin*, *39*(7), 3–6.

Christopherson, K. M. (2011). Hardware or wetware: What are the possible interactions of pedagogy and technology in the classroom? *Teaching of Psychology*, *38*(4), 288-292.

Ciraj, A.M., Vinod, P., & Ramnarayan, K. (2010). Enhancing active learning in microbiology through case based learning: Experiences from an Indian medical school. *Indian Journal of Pathology and Microbiology*, 53(4), 729-734.

Cleary, A. M. (2008). Using wireless response systems to replicate behavioral research findings in the classroom. *Teaching of Psychology*, *35*(1), 42-44.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159.

Corpuz, E.D., Corpuz, A.A., Rosalez, R. (2010). The use of web-based classroom interaction system in introductory physics classes. *AIP Conference Proceedings*, *1289*(1), 109-112.

Creswell, J.W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches.* Thousand Oaks, CA: Sage.

Creswell, J.W. (2015). *A concise introduction to mixed methods research.*. Thousand Oaks, CA: Sage.

Creswell, J.W., Plano Clark, V. (2011). *Designing and conducting mixed methods research.* Thousand Oaks, CA: Sage.

Crews, T.B., Ducate, L., Rathel, J.M., Heid, K., & Bishoff, S.T. (2011). Clicker in the classroom: Transforming students into active learners. *ECAR Research Bulletin, 9.* Retrieved from http://its.unl.edu/pdfs/Educause%20SRS%20success.pdf

Crouch, C. H., Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, *69*(9), 970-977.

Dallaire, D. H. (2011). Effective use of personal response “clicker” systems in psychology courses. *Teaching of Psychology*, *38*(3), 199-204.

Denzin, N.K. (1978). *Sociological methods.* New York, NY: McGraw-Hill.

Donovan, W. (2008). An electronic response system and Conceptests in general chemistry courses. *The Journal of Computers in Mathematics and Science Teaching*, *27*(4), 369-389.

Doucet, M., Vrins, A., & Harvey, D. (2009). Effect of using an audience response system on learning environment, motivation and long-term retention, during case-discussions in a large group of undergraduate veterinary clinical pharmacology students. *Medical Teacher*, *31*(12), e570-e579.

Dufresne, R. J., & et al. (1996). Classtalk: A classroom communication system for active

learning. *Journal of Computing in Higher Education, 7*(2), 3-47.
Eastman, J. K., Iyer, R., & Eastman, K. L. (2011). Business students' perceptions, attitudes, and satisfaction with interactive technology: An exploratory study. *Journal of Education for Business, 86*(1), 36-43.
Edmonds, C. T., & Edmonds, T. P. (2008). An empirical investigation of the effects of SRS technology on introductory managerial accounting students. *Issues in Accounting Education*, *23*(3), 421-434.
Elby, A. (2001). Helping physics students learn how to learn. *American Journal of Physics, 69*(S1), pp. S54-S64.
Elicker, J. D., & McConnell, N. L. (2011). Interactive learning in the classroom: Is student response method related to performance? *Teaching of Psychology, 38*(3), 147-150.
Fagen, A. P., Crouch, C. H., & Mazur, E. (2002). Peer instruction: Results from a range of classrooms. *The Physics Teacher, 40*, 206-209.
Field, A. (2009). *Discovering statistics using SPSS*. London, England: SAGE.
Fies, C., & Marshall, J. (2008). The C3 framework: Evaluating classroom response system interactions in university classrooms. *Journal of Science Education & Technology, 17*(5), 483-499.
FitzPatrick, K. A., Finn, K. E., & Campisi, J. (2011). Effect of personal response systems on student perception and academic performance in courses in a health sciences curriculum. *Advances in Physiology Education, 35*(3), 280-289.
Gentry, D. B. (2007). Using audience response systems in FCS. *Journal of Family and Consumer Sciences*, *99*(2), 42-44.
Ghosh, S., & Renna, F. (2009). Using electronic response systems in economics classes. *Journal of Economic Education, 40*(4), 354-365.
Gormley-heenan, C., & McCartan, K. (2009). Making it matter: Teaching and learning in political science using an audience response system. *European Political Science: EPS, 8*(3), 379-391.
Gravetter, F.J., & Wallnau, L.B. (2014). *Essentials of statistics for the behavioral sciences* (8th ed.). Belmont, CA: Wadsworth.
Gray, K., & Steer, D. N. (2012). Personal response systems and learning: It is the oedagogy that matters, Not the Technology. *Journal of College Science Teaching, 41*(5), 80-88.
Green, J. (2006). Toward a methodology of mixed methods social inquiry. *Research in the Schools. 13*(1), pp. 93-99.
Greene, J. C., & Caracelli, V. J. (2003). Making paradigmatic sense of mixed-method practice. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 91-110). Thousand Oaks, CA: Sage.
Hatch, J.A. (2002). Doing qualitative research in education settings. Albany, NY: State University of New York Press.
Hattie, J. (2008). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. New York, NY: Routledge.
Healy, M., & Perry, C. (2000). Comprehensive criteria to judge validity and reliability of qualitative research within the realism paradigm. *Qualitative Market Research, 3* (3), 118-126.
Hoekstra, A., & Mollborn, S. (2012). How clicker use facilitates existing pedagogical

practices in higher education: data from interdisciplinary research on student response systems. *Learning, Media and Technology, 37*(3), 303.

Hunter, A., Rulfs, J., Caron, J. M., & Buckholt, M. A. (2010). Using a classroom response system for real-time data display and analysis in introductory biology labs. *Journal of College Science Teaching, 40*(2), 19-25.

Hunter A., Revell, S. M., & McCurry, M. K. (2010). Engaging millennial learners: Effectiveness of personal response system technology with nursing students in small and large classrooms. *Journal of Nursing Education, 49*(5), 272-275.

d'Inverno, R., Davis, H. & White, S. (2003). Using a personal response system for promoting student interaction. *Teaching Mathematics and its Applications, 22*(4), 163-166.

James, M. C. (2006). The effect of grading incentive on student discourse in peer instruction. *American Journal of Physics, 74*(8), 689-691.

James, M. C., & Willoughby, S. (2011). Listening to student conversations during clicker questions: What you have not heard might surprise you! *American Journal of Physics, 79*(1), 123-132.

Johnson, R. B., & Christensen, L. B. (2004). Educational research: Quantitative, qualitative, and mixed approaches. Boston, MA: Allyn and Bacon.

Johnson, R.B., & Onwuegbuzie, A.J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14-26.

Kang, H., Lundeberg, M., Bjorn, W., delMas, R. & Herried, C.F. (2011). Gender differences in student performance in large lecture classrooms using personal response systems ('clickers') with narrative case studies. *Learning, Media and Technology, 37*(1), 53-76.

Karaman, S. (2011). Effect of audience response systems on student achievement and long-term retention. *Social Behavior & Personality: An International Journal, 39*(10), 1431-1440.

Keller, C., Finkelstein, N., Perkins, K., Pollock, S., Turpen, C., & Dubson, M. (2007). Research-based practices for effective clicker use. *AIP Conference Proceedings, 951*(1), 128-131.

Kennedy, G. E., & Cutts, Q. I. (2005). The association between students' use of an electronic voting system and their learning outcomes. *Journal of Computer Assisted Learning, 21*(4), 260-268.

Keough, S. M. (2012). Clickers in the classroom: A review and a replication. *Journal of Management Education, 36*(6), 822-847.

Kim, K., Sharma, P., Land, S., & Furlong, K. (2013). Effects of active learning on enhancing student critical thinking in an undergraduate general science course. *Innovative Higher Education, 38*(3), 223-235.

King, D. B., & Joshi, S. (2008). Gender differences in the use and effectiveness of personal response devices. *Journal of Science Education and Technology, 17*(6), 544-552.

King, D. B. (2011). Using clickers to identify the muddiest points in large chemistry classes. *Journal of Chemical Education, 88*(11), 1485-1488.

Kolikant, Y. B.-D., Drane, D., & Calkins, S. (2010). 'Clickers' as catalysts for transformation of teachers. *College Teaching, 58*(4), 127-135.

Krueger, R. A. (2000). Focus groups: A practical guide for applied research (3rd ed.).

Thousand Oaks, CA: Sage.
Lasry, N. (2008). Clickers or flashcards: Is there really a difference? *The Physics Teacher*, *46*, 242-244.
Lasry, N., Mazur, E., & Watkins, J. (2008). Peer instruction: From Harvard to the two-year college. *American Journal of Physics. 76*(11), pp. 1066-1069.
Lee, A., Lin, D., Reay, N. W., & Lei, B. (2011). Single-concept clicker question sequences. *Physics Teacher, 49*(6), 385-389.
Lincoln, D. J. (2008). Teaching with clickers in the large-size principles of marketing class. *Marketing Education Review*, *18*(1), 39-45.
Lincoln, Y.S., & Guba, E.G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
Lundeberg, M., Kang, H., Wolter, B., delMas, R., Armstrong, N., Borsari, B., Hagley, R. (2011). Context matters: increasing understanding with interactive Clicker Case studies. *Educational Technology Research & Development*, *59*(5), 645-671.
Madsen, A., McKagan, & Sayre, E. (2015). How physics instruction impacts students' beliefs about learning physics: A meta-analysis of 24 studies. *Phyical Review Special Topics - Physics Education Research. 11*(1), pp. 1-19.
Mayer, R.E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., Bulger, M., Campbell, J., Knight, A. & Zhang, H. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology*, *34*(1), 51-57.
Mazur, E. (1997). Peer instruction: A user's manual. Upper Saddle River, NJ: Prentice Hall.
McDermott, L.C., Shaffer, P.S., & the Physics Education Group at the University of Washington (2002). *Tutorials in introductory physics.* Upper Saddle River, NJ : Prentice Hall.
Micheletto, M. J. (2011). Using audience response systems to encourage student engagement and reflection on ethical orientation and behavior. *Contemporary Issues in Education Research*, *4*(10), 9-17.
Mordhorst, S. (2010). *Student Response Systems: The Impact on Nursing Student Engagement and Classroom Participation.* (M.S. 1504596), South Dakota State University, United States -- South Dakota. Retrieved from http://excelsior.sdstate.edu/login?url=http://search.proquest.com/docview/904412049?accountid=28594 Dissertations & Theses @ South Dakota State University; ProQuest Dissertations & Theses (PQDT) database.
Morling, B., McAuliffe, M., Cohen, L., & DiLorenzo, T. M. (2008). Efficacy of personal response systems ("clickers") in large, introductory psychology classes. *Teaching of Psychology*, *35*(1), 45-50.
Morote, E.-S., & Pritchard, D. E. (2009). What course elements correlate with improvement on tests in introductory Newtonian mechanics? *American Journal of Physics*, *77*(8), 746-753.
Morse, J. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research*, *40*, 120-123.
Nicol, D. J., & Boyle, J. T. (2003). Peer instruction versus class-wide discussion in large classes: A comparison of two interaction methods in the wired classroom. *Studies in Higher Education*, *28*(4), 457.
Oliver-Hoyo, M. T., Allen, D., Hunt, W. F., Hutson, J., & Pitts, A. (2004). Effects of an

active learning environment: Teaching innovations at a research I institution. *Journal of Chemical Education, 81*(3), 441-448.

Patton, M.Q. (1999). Enhancing the quality and credibility of qualitative analysis. *HSR: Health services research. 34*(5) Part II. pp. 1189-1208.

Perkins, K. K., & Turpen, C. (2009). Student perspectives on using clickers in upper-division physics courses. *AIP Conference Proceedings, 1179*(1), 225-228.

Porter, A. G., & Tousman, S. (2010). Evaluating the effect of interactive audience response systems on the perceived learning experience of nursing students. *Journal of Nursing Education, 49*(9), 523-527.

Pradhan, A., Sparano, D., & Ananth, C. V. (2005). The influence of an audience response system on knowledge retention: An application to resident education. *American Journal of Obstetrics & Gynecology, 193*(5), 1827-1830.

Prather, E. E., Rudolph, A. L., Brissenden, G., & Schlingman, W. M. (2009). A national study assessing the teaching and learning of introductory astronomy. Part I. The effect of interactive instruction. *American Journal of Physics, 77*(4), 320-330.

Premuroso, R. F., Tong, L., & Beed, T. K. (2011). Does using clickers in the classroom matter to student performance and satisfaction when taking the introductory financial accounting course? *Issues in Accounting Education, 26*(4), 701-723.

Price, E., De Leone, C., & Lasry, N. (2010). Comparing educational tools using activity theory: Clickers and flashcards. *AIP Conference Proceedings, 1289*(1), 265-268.

Quinn, A. (2010). An exploratory study of opinions on clickers and class participation from students of human behavior in the social environment. *Journal of Human Behavior in the Social Environment, 20*(6), 721-731.

Ray, d. I., Hugh, D., & Su, W. (2003). Using a personal response system for promoting student interaction. *Teaching Mathematics & its Applications, 22*(4), 163-169.

Redish, E. F., Saul, J. M., & Steinberg, R. N. (2000). Student expectations in introductory physics. *American Journal of Physics, 66,* 212-224.

Richardson, C. T., & O'Shea, B. W. (2013). Assessing gender differences in response system questions for an introductory physics course. *American Journal of Physics, 81*(3), 231-236.

Roxas, R.M., Carreon-Monterola, S.L., Monterola, C. (2010). Seating arrangement, group composition, and competition-driven interaction: Effects on students' performance in physics. *AIP Conference Proceedings, 1263*(1), 155-158.

Rudolf, A.L., Gonzaga, V., Prather, E., Brissenden, G., Consiglio, D. (2010). A national study assessing the teaching and learning of introductory astronomy part II: The connection between student demographics and learning. *Astronomy Education Review, 9*(1).

Sadler, P. M., & Tai, R. H. (2001). Success in introductory college physics: The role of high school preparation. *Science Education, 85*(2), 111.

Salemi, M. K. (2009). Clickenomics: Using a classroom response system to increase student engagement in a large-enrollment Principles of Economics course. *Journal of Economic Education, 40*(4), 385-404.

Schmidt, B. (2011). Teaching engineering dynamics by use of peer instruction supported by an audience response system. *European Journal of Engineering Education, 36*(5), 413-423.

Schwartz, M. S., Sadler, P. M., Sonnert, G., & Tai, R. H. (2009). Depth versus breadth:

How content coverage in high school science courses relates to later success in college science coursework. *Science Education, 93*(5), 798-826.

Seidman, I., (2013). *Interviewing as qualitative research: A guide for researchers in education and the social sciences*. New York, NY: Teachers College Press.

Sevian, H., & Robinson, W. E. (2011). Clickers promote learning in all kinds of classes-small and large, graduate and undergraduate, lecture and lab. *Journal of College Science Teaching, 40*(3), 14-18.

Shaffer, D. M., & Collura, M. J. (2009). Evaluating the effectiveness of a personal response system in the classroom. *Teaching of Psychology, 36*(4), 273-277.

Shapiro, A. M., & Gordon, L. T. (2012). A controlled study of clicker-assisted memory enhancement in college classrooms. *Applied Cognitive Psychology, 26*(4), 635-643. doi: 10.1002/acp.2843

Skiba, D. J. (2006). Got large lecture hall classes? Use clickers. *Nursing Education Perspectives, 27*(5), 278-280.

Smith, M. K., Wood, W. B., Krauter, K., & Knight, J. K. (2011). Combining peer discussion with instructor explanation increases student learning from in-class concept questions. *CBE - Life Sciences Education, 10*(1), 55-63.

Sternberger, C. S. (2012). Interactive learning environment: Engaging students using clickers. *Nursing Education Perspectives, 33*(2), 121-124.

Stowell, J. R., & Nelson, J. M. (2007). Benefits of electronic audience response systems on student participation, learning, and emotion. *Teaching of Psychology, 34*(4), 253-258.

Stowell, J. R., Oldham, T., & Bennett, D. (2010). Using student response systems ("clickers") to combat conformity and shyness. *Teaching of Psychology, 37*(2), 135.

Tai, R.H., Sadler, P.M., Loerh, J.F. (2006). Influencing chemistry success through high school chemistry teaching. *The Science Education Review, 5*(4), 123-127.

Tashakkori, A. & Teddlie, C. (2003). *Handbook of mixed methods in social and behavioral research.* Thousand Oaks, CA: Sage.

Teddlie, C., & Tashakkori, A. (2009). *Foundation of mixed methods research*. Thousand Oaks, CA: Sage.

Trees, A. R., & Jackson, M. H. (2007). The learning environment in clicker classrooms: student processes of learning and involvement in large university-level courses using student response systems. *Learning, Media & Technology, 32*(1), 21-40.

Trochim, M.K. (2006). *Research methods knowledge base*. Mason, OH: Thomson Custom Publishing.

Vaughn, S., Schumm, J.S., & Sinagub, J. (1996). *Focus group interviews in education and psychology.* Thousand Oaks, CA: Sage.

Webb, A., & Carnaghan, C. (2008). Investigating the effects of group response systems on student satisfaction, learning and engagement in accounting education. *Accounting Education,* 22(3), 391-409.

Webb, N. M., & Mastergeorge, A. (2003). Promoting effective helping behavior in peer-directed groups. *International Journal of Educational Research, 39*(1/2), 73.

Welsh, A. J. (2012). Exploring undergraduates' perceptions of the use of active learning techniques in science lectures. *Journal of College Science Teaching, 42*(2), 80-87.

Wieman, C., & Perkins, K. (2005). Transforming physics education. *Physics Today,* *58*(11), 36-41.

Wilkinson, S. (2004). Focus group research. In D. Silverman (ed.), *Qualitative research: Theory, method, and practice* (pp. 177-199). Thousand Oaks, CA: Sage.

Willoughby, S. D., & Gustafson, E. (2009). Technology talks: Clickers and grading incentive in the large lecture hall. *American Journal of Physics, 77*(2), 180-183.

Wolter, B. H. K., Lundeberg, M. A., Kang, H., & Herreid, C. F. (2011). Students' perceptions of using personal response systems ("clickers") with cases in science. *Journal of College Science Teaching, 40*(4), 14-19.

Yourstone, S. A., Kraye, H. S., & Albaum, G. (2008). Classroom questioning with immediate electronic response: Do clickers improve learning? *Decision Sciences Journal of Innovative Education, 6*(1), 75-88.