Comparing University Student Performance In Online V. Face-To-Face Offerings Of The Same Course, And Investigating Student Perceptions Of Satisfaction In An Online Course

Kristin Davidson
University of Mississippi

Follow this and additional works at: https://egrove.olemiss.edu/etd

Recommended Citation
Davidson, Kristin, "Comparing University Student Performance In Online V. Face-To-Face Offerings Of The Same Course, And Investigating Student Perceptions Of Satisfaction In An Online Course" (2016). Electronic Theses and Dissertations. 1046.
https://egrove.olemiss.edu/etd/1046

This Dissertation is brought to you for free and open access by the Graduate School at eGrove. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of eGrove. For more information, please contact egrove@olemiss.edu.
COMPARING UNIVERSITY STUDENT PERFORMANCE IN ONLINE V. FACE-TO-
FACE OFFERINGS OF THE SAME COURSE, AND INVESTIGATING STUDENT
PERCEPTIONS OF SATISFACTION IN AN ONLINE COURSE

A Dissertation
presented in partial fulfillment of requirements
for the degree of Doctor of Philosophy
in Higher Education
The University of Mississippi

by
KRISTIN ELIZABETH DAVIDSON

August 2016
ABSTRACT

Online education in the United States has seen dramatic growth for the past decade, outpacing any other growth in higher education. The concurrent mixed-methods study that was conducted for this research used data from a survey geology course taught in both environments, online and traditional face-to-face. The quantitative research focused on comparing student performance in an online course relative to the same face-to-face course, while the qualitative research investigated how students described their experiences taking an online class.

Previous work in online education has been limited by relatively small sample sizes, conducting studies over just one semester, comparing dissimilar courses in one study, considering few of the STEM disciplines, and, of the limited studies with GPA as a covariate, using self-reported GPA rather than actual GPA. The quantitative analysis of this study compared student performance in online (N=171) and face-to-face (N=1266) environments using data from the same STEM class over five years, with actual GPA as the covariate. ANCOVAs were calculated, and results showed that, overall, students performed better in the face-to-face class than in the online class, and this difference was more pronounced with students whose GPAs were 3.0 and lower. OLS regression was also conducted to identify predictors contributing to student success in the online classroom – GPA, course load, and student credit hours were the only significant factors predicting online performance.

For the qualitative component of this study, issues related to student satisfaction were explored by conducting a focus group from four students enrolled in the online STEM course.
Themes emerging from the discussion included interaction, technology, self-regulated learning practices, convenience, and course structure, with interaction as the most prominent theme. These findings help to explain the quantitative findings of why students with higher GPAs perform better – they do so, in part, because they have frequent interaction with the content despite the negative impact of the distance-based environment. Research, such as this study, is important in that identifying effective pedagogy promotes learning, particularly when the learning is done at a distance such as the online environment.
DEDICATION

I dedicate this to my husband, my soul mate, and my best friend, Gregg. You have been such an inspiration to me, and it is because of your constant support and encouragement through this journey that I finally find myself at its completion.

I also dedicate this to my four children, Kelly, Michelle, Kevin, and Heidi. Much of your lives you have seen your mother pursuing her educational goals, enough so, that lifelong learning seems normal to you. All of you are my joy and I love you.
ACKNOWLEDGEMENTS

There are many to acknowledge that have helped to make this research possible. Thank you first to Amanda Patterson and Cathy Grace. Without your input and willingness to share your data, I could not have done this study. You have been so enthusiastic in your support.

To Tiffany Ward, who was so encouraging to me when we met to discuss if this research was even possible. I was concerned I could not do a longitudinal study, getting the necessary snapshots of data, until I talked to you. Every obstacle I thought I might encounter, you responded to each one of them with, “No problem”. I remember leaving our meeting stunned and exhilarated, finally believing this whole process was really going to happen.

I thank my advisor, Dr. Lori Wolff, the first to introduce to me statistics in such a way that it made so much sense. From day one in your introductory statistics class, the first class I took in my doctoral program, I knew this was the path I wanted to take in terms of my research. Also, many “shout outs” for turning around my manuscript drafts so quickly. Your input has been invaluable and helped me stay-the-course in completing this research.

To my dissertation committee, Drs. Whitney Webb, Doug Davis and Qiang “Andy” Cheng. The classes I took from each of you was immensely helpful. When I started my Ph.D., I had not intended to do a mixed-methods study, but have since come to believe it has been a vital piece to understanding the numbers. At times, I felt I might be “biting off more than I could chew” but have found your input to help keep me focused and on track.
To my coders, Pari Bhatt and Joella Vaughnn. Both of you gave great input regarding my focus group. Joella, your help in some of the protocol, drawing from your own experiences in doing focus groups, gave me the confidence and foresight such that it all went surprisingly smoothly.

I also want to acknowledge Dr. Hsien-Yuan “Mark” Hsu. When I found myself stumped by my data, you willingly spent time with me to discuss a strategy. I was going down a path that was completely unnecessary and you helped me “see the light”, so to speak.

Thank you to my sister, Lisa Vaughnn. I cannot even begin to count the hours that we discussed this research. As mind-numbing as the details can be at times, you were always so enthusiastic and interested in it all.

To my parents, Bob and Carolyn Zolg. My entire life, you always made me believe that my goals could be accomplished. You have supported me and loved me in ways that words cannot begin to express. I will forever be grateful for the role-models you have been to me.

To my children, Kelly, Michelle, Kevin and Heidi. You four have had to endure the many long hours that your mother was devoting to her school. I do not ever remember you begrudging me when I had to “go work on the computer” or the late night dinners because I was in class.

To my husband Gregg. We did this journey of a doctorate 25 years ago for you and now find ourselves doing the same for me, even though I used to say I was not going to devote that kind of time to a doctorate like you did. You, most of all, have had to go through this process with me daily, and you so patiently and encouragingly did. Thank you for believing in me, challenging me, and supporting me through it all.
Finally, to my Lord and Savior, Jesus. It is through my faith in him that I have completed not only this journey, but all that I have accomplished.
# TABLE OF CONTENTS

ABSTRACT ................................................................................................................................. ii

DEDICATION ........................................................................................................................... iv

ACKNOWLEDGEMENTS .......................................................................................................... v

LIST OF TABLES .................................................................................................................... xii

LIST OF FIGURES .................................................................................................................. xiii

CHAPTER I: INTRODUCTION ................................................................................................. 1

Purpose Statement ..................................................................................................................... 4

Qualitative Research Question ............................................................................................... 5

Quantitative Research Question ............................................................................................. 5

Hypotheses ............................................................................................................................... 5

Significance of Study ................................................................................................................. 6

Definition of Terms .................................................................................................................. 8

Limitations and Delimitations ................................................................................................. 9

Summary .................................................................................................................................. 10

Overview of the Remainder of the Study ................................................................................. 10

CHAPTER II: REVIEW OF THE LITERATURE ...................................................................... 12

Learning Achievement ............................................................................................................ 13

    Learning Assessment .......................................................................................................... 13

    Self-Regulated Learning Model ......................................................................................... 14

    Transactional Distance Theory ......................................................................................... 16
Hypothesis One, Hypothesis Two, and Hypothesis Three ................................................................. 60
MatchIt Package in R .................................................................................................................. 60
ANCOVA Assumptions .............................................................................................................. 64
Robust ANCOVA Using R ...................................................................................................... 66
  Running R’s ancova and ancboot functions .............................................................................. 68
  Interpreting R’s Robust ANCOVA for Geol 105 Final Scores ................................................. 71
  R’s Robust ANCOVA for Geol 105 for Test 1 and Test 2 .................................................... 74
Hypothesis Four .......................................................................................................................... 76
  OLS Regression Assumptions ............................................................................................... 77
  SPSS Multiple Regression ...................................................................................................... 81
Summary ........................................................................................................................................ 86
CHAPTER V: QUALITATIVE FINDINGS ......................................................................................... 88
  Subject Selection .................................................................................................................... 88
  Planning and Conducting the Focus Group ........................................................................... 90
  Emerging Themes ................................................................................................................... 93
    Interaction ............................................................................................................................. 93
      Student-to-Instructor Interaction ....................................................................................... 93
      Student-to-Student Interaction ......................................................................................... 94
      Student-to-Content ............................................................................................................. 95
  Technology .............................................................................................................................. 96
  Self-Regulated Learning Practices ......................................................................................... 99
  Convenience ............................................................................................................................. 99
  Course Structure .................................................................................................................... 100
CHAPTER VI: DISCUSSION OF QUANTITATIVE AND QUALITATIVE RESULTS

Quantitative Results

Hypothesis One, Hypothesis Two and Hypothesis Three Conclusions

Hypothesis Four Conclusion

Using the OLS Regression Equation in Consideration of the Final Scores ANCOVA

Summary of the Quantitative Results

Qualitative Findings

Summary of Qualitative Findings

Future Work

Conclusion

LIST OF REFERENCES

LIST OF APPENDICES

APPENDIX A: EMAIL CORRESPONDENCE

APPENDIX B: FOCUS GROUP SEMI-STRUCTURED INTERVIEW QUESTIONS

APPENDIX C: R OUTPUT FOR ROBUST ANCOVA AND ROBUST BOOTSTRAP ANCOVA

VITA
LIST OF TABLES

1. Focus Group Semi-Structured Interview Questions .............................................................57
2. Percentage, Actual Sample Size, and Targeted Sample Size by Grade for the Face-to-Face Geol 105 Data ........................................................................................................................................62
3. Geol 105 Final Average, Exam Averages, and GPA Average by Grade from Complete Data Set ........................................................................................................................................63
4. Geol 105 Final Average, Exam Averages, and GPA Average by Grade from the MatchIt Data Set ........................................................................................................................................63
5. Course Environment Sample Sizes for Each Year of Data ......................................................73
6. Summary of GPAs Used in R’s Robust ANCOVA for Geol 105 Final Scores ......................74
7. Summary of GPAs Used in R’s Robust ANCOVA for Geol 105 Test 1 ...............................75
8. Summary of GPAs Used in R’s Robust ANCOVA for Geol 105 Test 2 ...............................76
9. Predictor Variables Used in Each Stage of Regression Analysis ...........................................81
10. Summary of the SPSS Coefficients Output, including only the necessary information for the OLS equation ......................................................................................................................84
11. Summary of the SPSS Coefficients Output by Year .............................................................85
12. Face-to-Face SPSS OLS Regression Results .........................................................................104
13. Comparison of Predicted Final Score in Geol 105 by Course Environment ......................105
LIST OF FIGURES

1. Moore’s Theory of Transactional Distance ................................................................. 17
2. Moore’s Dialogue Framework ...................................................................................... 17
3. Moore’s Transactional Distance Theory with the added tenet of technology .......... 19
4. SPSS ANOVA output of GPA and Course Environment .............................................. 65
5. R output for robust *ancova* and *ancboot* functions .............................................. 69
6. R plot of the final scores to GPA in the online and face-to-face data ......................... 70
7. Standardized residuals of Geol 105 final scores to standardized predicted values .... 78
8. SPSS Model Summary Output ..................................................................................... 79
9. SPSS excluded variables output for linear regression ................................................. 80
10. SPSS coefficients output .......................................................................................... 83
CHAPTER I
INTRODUCTION

Developing technologies in the last decade have led to a rapidly growing movement in online classes. Today’s students are accustomed to the use of technology in almost every aspect of their lives, and it follows that having the internet easily accessible – computers, tablets, and smart phones – should result in the incorporation of these and other emerging technologies in the classroom. In addition, online courses have evolved, from synchronous semester courses, to asynchronous independent study courses, to non-credit bearing massive open online courses (MOOCs) (Pence, 2012). In spite of increasing online course offerings, however, debate continues on the role online courses should play in higher education. In particular, questions remain on the effectiveness of the mode – online versus traditional face-to-face – in terms of student performance as well as student perception and satisfaction (Zemsky & Massy, 2004). Furthermore, college and university administrators have expressed concern of low retention rates in online courses (Allen & Seaman, 2015). Although national data is not differentiated by classroom environment, research estimates that as many as fifty percent of students drop out of online courses (Betts, 2008).

The Sloan Consortium (Sloan-C) has conducted extensive studies, beginning in 2003, related to general issues in online learning (Allen & Seaman, 2013). The results from questions on the College Board’s Annual Survey of Colleges pertinent to the online and traditional learning debate are published in the Sloan-C report every two years. This survey is administered to several thousand institutions with responses collected from college and university administrators.
In 2012, the Sloan-C Report published that 6.7 million students have taken at least one online class (Allen & Seaman, 2013). Moreover, online course growth has far outpaced any other growth in higher education, with many consecutive years of double digit growth rates in the last decade (Allen & Seaman, 2010). The online issue is so pertinent in today’s higher education institutions that 70% of college administrators responding to the 2013 Annual Survey of Colleges stated that online education was a critical component to their institutions’ long-term strategic planning (Allen & Seaman, 2013). Consistently over the past decade, the Sloan-C Reports have suggested that student demand for online course offerings continues to grow significantly with no foreseeable decline.

Online courses serve over 30% of the student population (Allen & Seaman, 2010), which emphasizes the importance of understanding the impact an online environment has on student satisfaction and student achievement. Researchers, however, have observed mixed results and still seek to answer the ongoing debate of student success in online courses as compared to traditional courses. Studies in technical courses such as statistics and geographical information systems (GIS), for example, have shown that students in online courses achieve the same, if not better, than those in traditional classrooms (Detwiler, 2008; Summers, Waigandt, & Whittaker, 2005). Other studies, however, contradict such findings. Edmonds’ (2006) study analyzed performance in a general psychology course, determining that online student performance significantly lagged those students in the traditional face-to-face delivery of the same course. In yet another study, conducted by Atchley, Wingenbach, and Akers (2013), using five years of archival data from Texas public universities, researchers compared data from students enrolled in traditional and online courses across fourteen disciplines. Similar to Edmonds’ (2006) study,
researchers determined students in traditional courses decidedly outperformed those in the online courses, notwithstanding the course discipline.

Not surprising, an added element to the debate regarding the value of online courses is the serious concerns of faculty regarding the effectiveness of online courses. Studies surveying professors of higher education institutions have expressed uncertainty that students in online courses have gained sufficient knowledge of the material as outlined in the course objectives (Cho & Berge, 2002; Maguire, 2005). Specifically, faculty members have stated that it is difficult to determine whether or not students are actually doing the work, and, if they are, there is doubt that those students successfully completing online courses have achieved the same learning outcomes as those in a traditional setting (Haber & Mills, 2008).

While the growth rate of students taking online courses in recent years has slightly declined, the overall number of students has continued to climb (Allen & Seaman, 2013). As the trend of student demand for online class offerings remains strong, understanding student perceptions and factors influencing student satisfaction in regard to online technology can provide valuable insight in how to best design and integrate such courses in higher education. Although perception is the cognitive process of deriving knowledge, it is more often thought of as simply a person’s reality – perception is one thinking something to be true and therefore it must be true (Otter et al., 2013). Accordingly, faculty members developing online technologies should consider student satisfaction and perception to positively impact a student’s overall experience in higher education. In Bolliger and Wasilik’s (2012) study assessing student perceptions and satisfaction levels of online courses, results showed that increased satisfaction affected higher retention rates and improved course performance. In particular, perception and satisfaction were sufficiently “influenced by personal attitudes, expectations, experiences, and
accomplishments in a specific setting and learning environment” (Bolliger & Wasilik, 2012, p. 154). Thus, understanding what factors contribute to student perception and satisfaction, specifically in online courses where students are not in the physical presence of a classroom with an instructor and other students, not only serves to enhance a students’ overall experience, but also contributes to both higher retention and achievement rates.

**Purpose Statement**

The purpose of this concurrent mixed-methods study was to identify those characteristics, including course environment that contributed to student learning, as well as those perceptions that influenced student satisfaction in the classroom. In this type of mixed methods study, data is collected in parallel, analyzed separately, and merged when interpreting the results. Scores from an introductory survey geology course, taught by instructors experienced in teaching in both types of environments – online and traditional face-to-face – were used to compare student performance by classroom environment. In order to better understand student performance, a focus group was conducted in which students described their experiences in the online environment and what perceptions they felt influenced their overall satisfaction in the course.

This course was purposefully chosen because it was not only taken by students across all the colleges and schools at the university to satisfy their general science elective requirement for their degrees, but it was also required as one of the core courses for Geology majors.

Furthermore, as the investigator, I personally knew these instructors and was familiar with the Geology department survey courses, including the course chosen for this study, which was first introduced by the department in the Spring 2010 semester. After having discussed the details of the research problem that was to be investigated, both instructors, from the online and traditional face-to-face geology courses, agreed to provide student data from the five semesters
they each taught the course, which, after approval from my dissertation committee approval and the Institutional Review Board (IRB Protocol #16x-201exempt under 45 CFR 46.101(b)(#2, 4)), made this study possible. In general, for a research study using this type of data, owners of the data and supervising administrators must first agree to release the data. Once agreement is given, all identifier variables, such as names and student identification numbers, must be removed from the data since it is protected under The Family Educational Rights and Privacy Act of 1974 (FERPA), which does not allow the release of educational records such that a student can be identified (http://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html). As long as data is released and complies with FERPA, such a study could be replicated.

**Qualitative Research Question**

How do students describe their level of satisfaction in online classes?

1. What perceptions contribute to student satisfaction in online classes?

2. How do students describe their self-regulating learning practices that contribute to student success in online classes?

3. What do students believe should be implemented to improve the delivery of online classes?

4. How do students’ perceptions and expectations of online classes differ from traditional, face-to-face, classes?

**Quantitative Research Question**

Does student performance differ by classroom environment: online versus traditional face-to-face?

**Hypotheses**
Hypothesis One: There is no significant difference in the mean score of test one by type of classroom environment (online or traditional) when controlling for incoming GPA.

Hypothesis Two: There is no significant difference in the mean score of test two by type of classroom environment (online or traditional) when controlling for incoming GPA.

Hypothesis Three: There is no significant difference in the mean score of the final grade by type of classroom environment (online or traditional) when controlling for incoming GPA.

Hypothesis Four: There is no significant relationship in the mean score of the final grade and a group of predictor variables, including GPA, prior online experience, major as categorized as STEM and non-STEM, major as categorized as Geology and non-Geology, number of completed student credit hours, and current course load (number of credit hours for the semester).

Significance of Study

Technology is widespread in today’s society, which has precipitated dramatic growth in online education. It is therefore important to further investigate the effectiveness of online courses in order to gain insight into how such emerging technologies’ impact successful learning practices. A review of the literature showed that understanding the factors that have contributed to success and perceptions that have shaped satisfaction in the online classroom are still uncertain. The ongoing debate of ascertaining the effectiveness of the online environment can best be answered by conducting studies that compare the two modes of delivery, online and traditional face-to-face. The intent of this study was to contribute to answers regarding the persistent questions behind traditional face-to-face and online environments by considering both student performance and student satisfaction.
In addition to contributing to current research by analyzing five years of longitudinal data from a survey geology course regarding the four hypotheses, this study also considered those disciplines in science, technology, engineering, and math (STEM). Most studies conducted to date have not specifically considered STEM in the online environment. The one study done using STEM courses was at the community college level and only investigated online STEM courses required for a degree versus online STEM courses satisfying an elective (Wladis, Hachey, & Conway, 2014). Having used data from an elective science course at a research university, this study looked specifically at how students performed in a STEM class relative to their major. Data from this study served to identify whether students in STEM disciplines were better suited to take online STEM classes than those in non-STEM disciplines. Specifically, this study investigated whether students taking the STEM online class who had also declared STEM majors achieved higher than those who had declared non-STEM majors. Moreover, this study also considered whether students who had specifically declared Geology as their major achieved higher than all other majors. Results of this study have helped to contribute to this void in research in better understanding the STEM role in the online classroom.

Finally, few studies have considered student characteristics that influence both a student’s performance and satisfaction in web-based courses. This study investigated the online delivery mode in terms of what predictors, if any, influenced a student’s academic success, and what perceptions framed a student’s level of satisfaction. Such factors serve to aid faculty members in designing more pedagogically sound online courses. Moreover, knowing what characteristics impact student success also add to the current literature, revealing those student strategies and motivational practices leading to the greatest success in online learning. Currently, one of the biggest issues for many students is a failure to realize that the two modes of instruction often
require fundamentally different learning strategies. Specifically, students are accustomed to the frequent interaction with instructors and other classmates when in the face-to-face classroom. Many of these individuals only consider the convenience of remote attendance when enrolling in an online class, without concern of the necessary self-regulated and motivational practices required for successful completion of a course delivered in this mode. Furthermore, students who fall into this category often struggle to successfully complete online courses because of unanticipated obstacles, most notably the lack of interaction with the instructor and other students (Kranzow, 2013). By considering performance and satisfaction, online courses can be designed to help students quickly ascertain whether the online environment is suitable for them, and data from this study helped to identify some of those influences.

Definition of Terms

1. **Course outcomes.** Index for evaluating the quality of an academic program, comprised of learning achievement and student satisfaction (Wang, Shannon, & Ross, 2013).
2. **CAQDAS.** Computer-assisted qualitative data analysis software.
3. **Face-to-face (F2F) course.** A course in which all content is delivered orally or in writing (Allen & Seaman 2008).
5. **Geol 105.** Geology – Resources, a survey geology course covering natural resources, waste disposal, and climate change.
6. **GPA.** Grade point average.
7. **IREP.** Institutional Research, Effectiveness and Planning
8. **Learning achievement.** Cognitive variable of course outcomes (Wang et al., 2013).
9. **MOOC.** Massive open online course.
10. **OLS.** Ordinary Least Squares regression, also referred to as multiple regression.

11. **Online course.** A course in which at least 80% of the content is online (Allen & Seaman, 2008)

12. **PSM.** Propensity Score Matching.

13. **Self-regulated learning (SRL).** A students’ active and self-motivated behavior to perform academic tasks (Wang et al., 2013).

14. **Sloan-C.** Sloan Consortium.

15. **Student satisfaction.** Affective variable of course outcomes (Lee, 2014).

**Limitations and Delimitations**

One limitation of the quantitative component of this study was the evaluation of achievement from only one type of geology survey course. While few studies have considered university STEM courses (Wladis et al. 2014), choosing a particular STEM course, geology in this research, may not be generalizable to all similar types of STEM courses, particularly elective science courses.

Moreover, the instructors for each of the two delivery modes were different, albeit the same textbook and same test banks were used. Even though the instructors were veteran teachers, both having taught online and traditional classes for many years, this difference could have impacted the results of student achievement and satisfaction.

A second limitation was the disparate enrollment sizes of the two environments. The traditional face-to-face courses that were used in this study had almost three times the number of students than that of the online courses. Statistical methods, particularly regression analyses, are most reliable when the sample sizes that are compared are relatively close in number (Hinkle, Wiersma, & Jurs, 2003). Although this study employed a robust method that minimized the type
I statistical error when analyzing an unbalanced sample size – students in each delivery mode for this study – the statistical results may not be generalizable to similar online survey geology courses.

Another limitation related to the disparate time period in which the quantitative and qualitative data were collected. The concurrent mixed-methods approach of this study was for the qualitative data that was collected to help broaden the perspective and understanding of the quantitative data results. The quantitative data was from 5-years of archival data. The qualitative data, however, was from just one semester. Although the students’ perceptions related to their own experiences, it may not have explained other students’ perceptions of the same online geology environment.

Finally, a limitation of the qualitative component of this study resulted from the small group of students the researcher selected to participate in the semi-structured interview. The intent of the focus group was to describe student perceptions of satisfaction in an online environment, including those learning strategies that were used during the semester. While qualitative research is subjective, it is important that it does not undermine understanding the general themes for improving learning strategies and satisfaction in the online environment (Patton, 2002). It is possible that those themes that emerged from the focus group may only be true for those participants in the group, not for a general understanding of student satisfaction in a similar online geology class.

Summary

Factors relating to student performance and overall satisfaction are important not only for academic institutions in their strategic planning, but also for instructors in their approach and preparation to teach a course in a technological format, as well as for students in their
expectations and plans for devising self-regulating learning strategies to be successful in online classes. Continued research can help to explore best practices leading to best outcomes, utilizing effective pedagogy that promotes learning in both the traditional as well as the online classroom environment.

**Overview of the Remainder of the Study**

The following chapters provide more in depth rationale of the study. Chapter two details the theoretical framework as well as a comprehensive review of the literature related to online learning issues, particularly student achievement and student satisfaction studies in such a learning environment. Chapter three discusses the methodology used to conduct the study, which includes specifics on the statistical analyses to evaluate the quantitative data and the interview structure of the focus group for the qualitative data. Chapters four and five detail the results and findings of the data, and chapter six is a discussion of the conclusions from the study.
CHAPTER II
REVIEW OF THE LITERATURE

This literature review investigates the role of technology in the classroom, specifically the online environment in how it compares to the traditional face-to-face environment. The framework for online learning is rooted in both the Self-Regulated Learning (SRL) model and Moore’s (1993) theory of transactional distance. Specifically, transactional distance considers how interactions among each of the individuals in the environment relate to the process of learning.

Two prominent themes have been associated with higher education online learning. The first addresses course outcomes related to those students enrolled in online classes relative to those in traditional face-to-face classes. The second considers key characteristics that contribute to higher student satisfaction rates, which have been shown to directly correlate to student success and retention rates.

This literature review begins by first examining learning achievement, followed by a discussion of both the SRL model and Moore’s (1993) transactional distance theory, each of which support the basis of this study. After discussion of these contexts, studies related to courses in the online environment are reviewed as well as studies comparing achievement in the online delivery mode relative to the traditional face-to-face delivery mode. Finally, an overview is given of literature studies on student satisfaction in online classes and the influence satisfaction has on student achievement.
Learning Achievement

Learning is the change in one’s thinking that shapes an individual’s general understanding (Entwistle & Peterson, 2004). Identifying predictors for the greatest success in student learning has continued to be an important issue in higher education, particularly for colleges and universities, whose mission is to teach and train the future generation (Thielens, 1977).

In studying memory, Craik and Lockhart (1972) coined the terms deep and surface in how it pertains to learning. Deep learning occurs when an individual is interested in the material, so much so the student is self-motivated to develop an extensive understanding of the subject. Alternately, surface learning is the result of coping with material to successfully pass the class, where the student typically invests minimally in the course objectives, most notably by memorizing the material with little regard to understanding the content. There is a place, however, for both. Saljo’s (1979) study of learning revealed that sometimes for one to expand understanding of certain subject material, memorization, or surface learning, may be necessary at some point in the process of achieving deep learning (e.g., mathematics, biology, and chemistry).

Learning Assessment. For faculty, assessments should be designed such that the content is promoting deep learning. Chickering and Gamson (1987) identified seven tenets of good practices in undergraduate education to help facilitate the deep learning process, regardless of the learning environment: (a) encourage contact between students and faculty, (b) develop reciprocity and cooperation among students, (c) encourage active learning, (d) give prompt feedback, (e) emphasize time on task, (f) communicate high expectations, and (g) respect diverse talents and ways of learning. Within these guidelines, Chickering and Gamson (1987) identified assessment as an important tool for ascertaining a student’s level of understanding, since
“knowing what you know and don’t know focuses learning” (p. 4). Assessments, they recommend, should be offered in a variety of ways, considering both theoretical and hands-on approaches. Additionally, instructors should be prompt in evaluating student performance so students know what they have learned and what they still need to learn.

Chickering and Ehrmann (1996) expanded on Chickering and Gamson (1987) by adding technology as a tool that should be used for enhancing instructional strategies. As Chickering and Ehrmann (1996) suggest, there is a place for technology in the classroom, but one that efficiently and effectively promotes learning, not one that is contrary to those principles indicative of good practices in undergraduate education.

In terms of assessment, Chickering and Ehrmann (1996) note that technology increases opportunity for learning not only in faster communication between the student and the instructor and the student and other students, but also in the types of assessments that can be offered, one’s that are “interactive, problem oriented, relevant to real-world issues, and that evoke student motivation” (Chickering & Ehrmann, 1996, p. 6). The authors also argue that it is important for students to identify those types of learning environments that serve them best, delivered through classroom lecture and the internet alike. It is through self-assessment that students can identify those practices necessary to be successful in the learning process.

**Self-Regulated Learning Model.** For college students, investing in deep learning throughout their college tenure is of great importance in order to attain long term personal and economic goals (Hannon, 2014). Indeed, understanding the nuances of deep learning begs the question, what strategies should students employ to engage in learning? Weinstein, Acee, and Jung (2011) describe learning strategies as behavioral goals an individual sets for himself/herself to acquire knowledge. Zimmerman and Martinez-Pons (1986) expanded this definition to
consider self-regulated learning strategies. Self-regulated learning is defined as those purposeful actions taken by an individual to acquire information or skills by setting personal goals and by self-evaluating achievement through observation, assessment and response relative to the degree of learning that has taken place – metacognitive, motivational, and behavioral responses to an individual’s progress in the learning process (Zimmerman & Martinez-Pons, 1986).

Pintrich (2004) developed a conceptual framework for evaluating self-regulated practices that college students employ in the learning process, which are based on four assumptions related to Zimmerman’s (2000) learning model – cognition, motivation, behavior, and context. To begin, cognition is the learner’s active and direct role in his/her learning. Motivation follows and is the learner’s ability to monitor his/her behavior in terms of what motivates him/her to learn. Motivation, however, is not to be mistaken for an individual always engaging in those behaviors that facilitate learning, but rather that he/she possesses the ability to monitor such strategies. Third is behavior, which is the learner’s ability to take action and establish and assess goals throughout the learning process. Finally, context is the direct relationship between the learner’s self-motivating practices and his/her achievement in the classroom.

In establishing these general assumptions related to the SRL model, Pintrich (2004) identified four phases a college student may experience in terms of motivation and learning outcomes. These phases are numbered one to four, although, Pintrich (2004) notes that they are not always sequential in practice. A student may not move in order from phase one to phase four, nor will a student necessarily enter all phases throughout the learning process.

Phase one of the SRL model involves the planning and forethought a student considers when establishing goals for succeeding in the learning process. Phase two occurs when a student monitors his/her progress in the learning process, including an awareness of how he/she is
performing in terms of meeting one’s goals. Phase three is control, or a student’s assessment and possible adaptation of learning strategies to be successful in his/her achievement. Last is a student’s reaction and reflection of how successful he/she has been in meeting the goals established at the beginning of the learning process.

In terms of online education, King, Harner, and Brown (2000) identified the necessity of student motivation to self-regulate his/her learning process, particularly because of the limitations specifically related to distance in such an environment as described by Moore’s (1993) theory of transactional distance. King et al. (2000) evaluated questionnaires from students at a four-year institution enrolled in distance education accounting and vaccinology courses (N = 113). Findings from their study determined it was critical a student possess motivational practices when learning online, even more so than those self-regulating practices students undertake when enrolled in traditional face-to-face courses. Therefore, it is important to consider the impact that distance has on a student’s ability to self-regulate deep learning. Moore’s (1993) theory of transactional distance specifically addresses the nuances to consider when the student and the classroom are physically separated.

**Transactional Distance Theory.** Moore’s (1993) theory of transactional distance considers the relationship between the teacher and the learners, particularly when there is physical distance between the instructor and the student. Moore’s premise is that pedagogical practices should be implemented to minimize the degree of *transactional distance* between students and instructors by maximizing positive interactions between students, instructor, and content (Offir, Lev, & Bezalel, 2006). Moore’s (1993) theory identifies three overarching tenets that impact transactional distance in learning, namely, dialogue, structure, and learner autonomy.
Each of these principles should be considered in the development of courses employing a distance learning environment.

**Figure 1.** Moore’s Theory of Transactional Distance shows a flow chart detailing the three factors that impact learning when physical distance separates the student and the location of the course.

Figure 2 outlines the first tenet, dialogue, which considers three types of positive interactions that occur in a learning environment – between students, between student and instructor, and between student and content (Ekwunife-Orakwue & Teng, 2014). Together, the instructor and each of the students has a purpose in the learning environment by offering contributions that help shape the knowledge of the community of learners (Moore, 1993).

**Figure 2.** Moore’s dialogue framework shows the interactions that occur when considering the dialogue tenet of Moore’s theory of transactional distance. A student will interact with other students, with the instructor, and with the content.
By considering the interplay of these three types of interactions in terms of dialogue, especially in regard to the role of students to instructors in such an environment, distance education courses can be designed to maximize dialogue in spite of the physical separation of the participants in the course.

The second tenet, structure, is the actual course management platform that is used to administer the course. It is through structure that student achievement can be assessed, evaluating “objectives, implementation procedures, and evaluation procedures of teaching programs” (Moore, 1980, p.21). The weaker the structure of the content delivery, the greater the need for dialogue between instructor and learner, in order to reduce the negative impact of transactional distance. If improving dialogue is not a practical solution, adjustments can be made to the structure to improve the learning environment, an “educational program’s responsiveness to the learner’s individual needs” (Moore, 1980, p. 21).

The last tenet, learner autonomy, considers students’ abilities to self-regulate their own learning – the SRL model – by establishing goals and assessing progress in a distance environment (Ustati & Hassan, 2013). Moore (1989) defined this interaction as “the process of intellectually interacting with the content that results in changes in the learner’s understanding, the learner’s perspective, and the cognitive structures of the learner’s mind” (p. 2). When a student regularly invests in comprehending new concepts, they acquire new knowledge, effecting tangible learning outcomes. In a learning environment where the instructor and learner are remote, to achieve these outcomes, the learner must have a high level of autonomy to overcome the lack of dialogue and structure that result from the transactional distance.

**Transactional Distance Theory and the Online Environment.** Although Moore’s theory was in regard to the traditional distance learning model, Bernard et al. (2009) considered
pedagogical features across several types of distance education platforms, including online courses (Abrami et al., 2011). The results of Bernard et al.’s (2009) study confirmed that Moore’s theory of transactional distance directly applies to online education, particularly the notion that student achievement and satisfaction positively relate to the degree of interaction among students, instructor, and content in a learning environment (Abrami et al., 2011). This complements Moore’s observation that teacher dialogue, content structure, and student autonomy are all influenced by transactional distance, including in the web-based delivery mode.

*Figure 3.* Moore’s Transactional Distance Theory with the added tenet of technology. In Moore’s theory of transactional distance, the dialogue tenet is expanded to include technology when considering courses offered in an online environment. Distance impacts a student’s learning not only in the student-to-instructor, student-to-student, and student-to-content interactions, but also the student-to-technology interaction when the course is offered in a web-based delivery mode.

Furthermore, Strachota (2003) posited an additional key factor to Moore’s transactional theory, specifically in terms of online education – the interaction between student and technology (Figure 3). This measure considers students’ technological self-efficacy, or the belief that they can master the technology required to be successful in an online class (Wang et al., 2013). To minimize the transactional distance in an online class, students must consider their abilities, or at
least their belief in their abilities, to use the technology required for online learning. When considering all four – student to student, student to instructor, student to content, and student to technology – online education can be designed to have the greatest success in not only engaging learners, but also in improving student achievement and satisfaction.

**Learning in an Online Environment as Compared to a Traditional Face-to-Face Environment.** Since 1990, the rate of online course enrollment has rapidly outpaced that of traditional enrollment, but the question remains whether learning achievement and student satisfaction in the online environment is as robust as that of traditional face-to-face (Wang et al., 2013). As Chickering and Ehrmann (1996) noted, technology in the classroom should promote learning. By comparing student achievement in both the online and traditional delivery modes, educators can gain a better understanding of how to implement technology such that it enhances student learning potential rather than one that is simply didactic, requiring only rote memorization.

An ongoing discussion point in online education relates to student learning outcomes. Are students learning as much in an online course as they would taking the same course in a traditional environment instead? One question from the Annual Survey of Colleges assessed administrators’ views of learning outcomes in online classes as compared to traditional classes. Seventy-five percent of public institution administrators and fifty-seven percent of private institution administrators felt online courses were as good, if not better, than traditional courses in terms of learning outcomes (Allen & Seaman, 2013). For faculty members, however, the numbers were not as optimistic. Although the percentage has slightly decreased over the last decade, nearly 25% of faculty still believe online courses are inferior and do not produce the same learning outcomes as those students taking traditional courses (Allen & Seaman, 2015).
Research regarding the success of student performance in online classes as compared to traditional classes continues to be debated. Typically, measures for performance are based on assignment grades, final grade in the course, or successful completion of the course. In their study, Driscoll and colleagues contributed to a better understanding of this issue by evaluating student performance in an introductory sociology course at a four-year institution (Driscoll, Jicha, Hunt, Tichavsky, & Thompson, 2012). For this study, conducted over three semesters with the same instructor, the research measured student performance by grades on a midterm course exam and a data analysis assignment (N=368). One aspect of the debate is whether it is just the learning environment that can be attributed to a student’s performance or if other factors have a significant impact as well. Therefore, Driscoll et al. (2012) considered several control variables in their research, including self-reported grade point average (GPA), time devoted to the course, current course load, number of hours worked at a paying job, current student classification (freshman, sophomore, junior, senior), number of completed online courses, and learning style. By introducing GPA as a control variable and using an ordinary least squares (OLS) regression, the results showed that there was no significant difference in online and traditional student performance in regard to either the exam or the analysis assignment. None of the other control variables had impact. Accordingly, this study demonstrates the importance of considering covariates in terms of student performance, particularly GPA, which was shown to have a significant statistical influence on the results if it had not been considered as a covariate in the calculations.

Frantzen (2013) suggested in his study at a southeastern university that the first strategy to consider in terms of online courses is “matching the technology with the course content and the mode of course delivery, and not the content with the technology” (p. 568). Consequently,
this study’s data collection was from three technology-driven assignments for a criminology course, using the students’ grades on the coursework as well as their final grade in the class (N=244). Results were calculated using an OLS regression while controlling for varying student characteristics such as GPA. Student performance was the same in the online classes as the traditional classes. Interesting findings from the results showed that students in the online classes did not turn in all of their technology assignments, and even with these missing assignments, their overall performance was still the same as those in the face-to-face classes. Also, the conclusions of this study confirmed the findings of Driscoll et al. (2012) in that GPA was determined to be the strongest predictor for student achievement, not course delivery mode.

In their research, looking at both student retention and learning outcomes at several community colleges in the northeast, Hachey, Wladis, and Conway (2012) evaluated whether a student’s previous online experience contributed to their success in future online courses (N=962). Using a binary regression analysis, researchers determined that previous online experience did affect a student’s performance in future online courses. Specifically, the data showed that students who had been successful in all their online classes were significantly likely to pass their future online classes.

Hachey, Wladis, and Conway (2014) later conducted a study, also using community college data, to identify whether a student’s prior success in an online course was the predictor alone for success in future online courses, or if other predictors such as pre-course GPA were more accurate (N=962). Using binary logistic regression, the data showed that students who had not previously taken an online course were more likely to succeed the stronger their GPA, confirming both the findings of Driscoll et al. (2012) and Frantzen (2013). The data for students who had previously taken at least one online course, however, GPA showed no bearing on their
success at all. Rather, their performance in their previous online courses alone served as the strongest predictor for success in future courses – they were more likely to succeed had they passed their previous online classes, and more likely to fail had they been unsuccessful in their previous online classes, either from dropping out or from failing the course.

Wladis, Hachey, and Conway (2014) further investigated online science, technology, engineering, and math (STEM) course outcomes at the community college level by determining whether or not online STEM course offerings had higher student success rates because they were core STEM classes as opposed to elective STEM classes, or if these courses were taken by students who showed certain characteristics that influenced their performance in the class, regardless of their major (N=3599). Again, using binary logistic regression, findings from the study showed that a student’s ability to succeed in an online STEM course was the result of characteristics they possessed rather than whether the course was either a fulfillment of their degree program or just an elective. Several important details were missing from this study, however, particularly what constitutes required versus elective. Specifically, for students claiming the class was elective, it is unknown whether these students were STEM or non-STEM majors. For some, those stating that the class was an elective, as opposed to required, may have made such a claim because of the small pool of STEM classes from which they could choose to fulfill this part of their degree program. Moreover, the study was at the community college level, and determining degree requirements was left ambiguous – was the STEM course credit-bearing for an Associate’s degree or was it also credit-bearing for a Baccalaureate degree, should the student transfer to a four-year institution?

In yet another study, conducted by Atchley et al. (2013), online course completion relative to face-to-face course completion and online course completion by discipline were
evaluated using archival data from Texas public universities (N=319,153). For the first part of their study, using a Goodness-Of-Fit Chi-square statistical method, the researchers’ results were opposite from that of Wladis et al. (2014) – students in traditional courses had significantly higher course completion rates than those in online courses. Unlike Wladis et al. (2014) and Driscoll et al. (2012), however, this study did not consider any factors, such as GPA, other than course completion. Based on previous studies, GPA is an important factor that should be considered. Without it, the results may be erroneously influenced by student aptitude – students are inclined to do better solely because they are a higher achieving student, not because of the environment in which they are taking a class.

Atchley et al. (2013) also used a Goodness-Of-Fit Chi-square statistical method to evaluate online course completion of fourteen distinct disciplines – accounting, agricultural education, computer information systems, English, finance, general business, health, human resource management, management, marketing, physical education, psychology, reading, and special education. The researchers’ results showed that students in certain disciplines, such as reading, were significantly more likely to complete an online course than those enrolled in other disciplines, such as finance. This finding suggests some disciplines are not as optimal for the online environment in terms of student retention.

In the second part of their study, Atchley et al. (2013) considered online and traditional courses that were taught by the same instructor, using the final grade as the measurement. The results of this study had an opposite conclusion of other researchers, emphasizing the ongoing disagreement of whether classroom environment matters in terms of delivery mode (Driscoll et al., 2012; Frantzen, 2014). Using a Goodness-Of-Fit Chi-square statistical method, the findings showed students in traditional course offerings outperformed students in online course offerings.
in terms of final grade. Also, regarding retention, students in traditional classes were more likely to complete a course rather than drop out. Unlike studies conducted by researchers such as Driscoll et al. (2012) and Frantzen (2013), however, control variables were not introduced to determine if these might have influenced student performance, explaining away the differences between students enrolled in traditional classes as compared to those enrolled in online classes.

Additionally, Xu and Jaggars (2014) investigated performance gaps of various student groups in community college online courses (N=116,830). Specifically, gap comparisons were evaluated regarding gender, ethnicity, age, and previous academic performance. For gender, the results showed a performance gap of women over men in both course persistence and better standardized course grade. Although this gap was true for both online and traditional lecture formats, it was significantly wider in online. For ethnicity, online course persistence was not noteworthy for any specific racial group. For standardized final grade, however, African American students were significantly more likely to earn a lower grade in an online course than any other ethnic group. For age, older students’ course persistence and standardized course grade were actually better in online classes rather than traditional classes. Older students’ overall performance in online courses was also better than that of younger students. Finally, previous academic performance was considered, whether or not a student entering community college was required to take any remedial courses. It was determined that students with weak academic preparation had a wider online performance gap than those with strong academic preparation.

At Baylor University, Carter and Emerson (2012) directed research using seven sections of a Principles of Economics course, three online and four traditional face-to-face (N=186 total). In this particular study, scores on a standardized Economics exam, Test of Understanding in
College Economics (TUCE), as well as overall course grade were used to determine learning outcomes. Over 90% of students enrolled in these sections were required to take the course for their degree program. Using OLS regression, and controlling for gender, age, ethnicity, professor-fixed effect, the researchers determined, as in previous studies, that there was no difference in achievement with regard to delivery mode, students in the online sections performed the same as those students enrolled in the face-to-face sections (Carter & Emerson, 2012). While Driscoll et al. (2012) and Frantzen (2013) did additional evaluation of the control variables to rationalize using GPA, no discussion was given in this study as for the justification of the covariates that were introduced.

In a General Psychology course at a small Ohio liberal arts college, Edmonds (2006) considered students enrolled in both online and traditional sections (N=175). Applying a MANOVA and controlling for high school GPA and SAT score, final letter grades for students in both environments were used as the basis for the study. As Atchley et al. (2013) also found in their study, students opting for the online delivery mode did not perform as well as those students in the face-to-face.

Various additional studies comparing online and traditional courses have been conducted, albeit with much smaller sample sizes. Detwiler’s (2008) study compared students in the online (N=19) and traditional (N=11) environments of a Geographical Information Systems (GIS) class. Because the course fulfilled the requirements of both a core course in a resident Geography Bachelor of Science degree program as well as a post-baccalaureate GIS certification program, the two populations of students significantly varied. The average age of those taking the course in the face-to-face environment toward their Bachelor of Science degree was 21, while the average age of those taking it for GIS certification was 40. Furthermore, some of the online
students included experienced programmers, a distinct advantage in a GIS course regardless of learning environment. Consequently, Detwiler (2008) disregarded these students in his study since an initial analysis of the data obviously skewed the results in favor of the online students outperforming the traditional students, an assessment influenced by a student’s prior programming skills rather than a determination of the learning environment’s impact on student achievement. Applying a Mann-Whitney U one-tailed test, the results of the study showed that those in the online environment performed the same if not slightly better than those in the face-to-face. Although analysis did not control for age, Detwiler (2008) concluded that those enrolled in the online course had “a greater level of motivation and maturity” (p. 143) as shown by not only their greater overall performance in the course but also their high level of self-regulated learning, spending “26% more time in their coursework than their on-campus counterparts” (p. 143). One additional factor not accounted for in this study, however, regards the incoming, or *apriori*, experience of the online students. The study only removed those students with prior programing skills from their sample set, but the researchers did not consider that those online students working in the industry were also coming into the course with previous experience and knowledge in GIS, a distinct advantage when trying to learn the software implementing these principles.

Summers et al. (2005) conducted a study of nursing students enrolled in a statistics course required for the Bachelor of Science in Nursing at a four-year institution (N=38). Of the 38 students enrolled, 17 chose to take the online version while 21 took the traditional face-to-face version. The study considered not only course outcomes, comparing exams scores in both environments, but also student satisfaction, assessing student evaluations. Because the results could also be skewed by some students having a prior understanding of statistics, the researchers
administered a pre-test and post-test on basic statistical knowledge. Using independent t-tests, the results from Summers et al. (2005) showed similar findings to previous studies in that of the two groups, online versus face-to-face, neither showed a significant difference in course outcomes given that the students began the course with the same level of understanding in statistics as shown by the pre-test and post-test results (Carter & Emerson, 2012; Detwiler, 2008; Driscoll et al., 2012; Frantzen, 2013; Hachey et al., 2014).

Jaggars, Edgecombe, and Stacey (2013b) conducted research comparing drop-out and failure rates of students in both online classes and traditional classes over three consecutive semesters at various community colleges. Moreover, this study considered additional information, looking at trends and possible performance gaps of various student populations. Although the sample size was quite small for the twenty three entry-level courses examined (N=46), their study considered five possible student groups in terms of retention and failure rates of online and traditional courses: overall starting enrollment of online and traditional courses, those in gatekeeper math and English courses, those in remedial math and English courses, those taking courses after successful completion of remedial math and English courses, and the overall performance gaps of different student demographics (Jaggars et al., 2013a). In all five cases, their research confirmed the conclusions of Atchley et al. (2013) – students in online classes consistently failed or dropped out more often than those students in similar face-to-face courses. One interesting discovery of this study was related to performance gaps. The researchers suggested that, in general, male students with low GPAs as well as students of African American ethnicity tended to perform lower than the general student population, and these differences were more pronounced when the classes were taken online. Of most significance, lower performing students, those with low GPAs, had a much steeper failure rate than those who were enrolled in
traditional classes instead. Similarly, African American students over White students had a sharper dropout rate when enrolled in online classes rather than traditional classes.

**Satisfaction and Learning Achievement**

Student satisfaction and perception is an integral part of a student’s overall experience in higher education. It is a student’s perception that affects both retention and performance as it is “influenced by personal attitudes, expectations, experiences, and accomplishments in a specific setting and learning environment” (Bolliger & Wasilik, 2012, p. 154). Moreover, when the course is taught at a distance, satisfaction is a key factor in not only course effectiveness but also student success in the learning process (Biner, Welsh, Barone, Summers, & Dean, 1997).

In the online environment, various predictors have been identified that contribute to student satisfaction: self-regulated learning strategies, technology self-efficacy, autonomy, and interaction (Artino & McCoach, 2008; Bolliger & Martindale, 2004; Reinhart & Schneider, 2001). Understanding how these student satisfaction factors contribute to a students’ overall experience can help to improve online courses, one that promotes student learning and strengthens student achievement rates (Reinhart & Schneider, 2001).

**Satisfaction in the Context of the Online Environment.** In their research, Wang et al. (2013) considered the question of how student characteristics and student satisfaction influenced learning outcomes in online courses. The model used for the study was purposefully based on models from previous research in order to further ascertain which factors positively influenced student outcomes and satisfaction. Specifically, the researchers wanted to evaluate whether race, gender, and educational level impacted a student’s motivation to self-regulate the time devoted to their learning and consequently achieve in a technological environment. The researchers considered four postulates in terms of how previous online experience was influenced by student
characteristics (race, gender, and educational level): self-regulated learning, technology self-efficacy, course satisfaction, and performance, where technology self-efficacy was defined as the belief that one can master the technology (Wang et al., 2013). Using a Goodness-Of-Fit Chi-square statistical method, the results of the study showed that students with previous online experience had more effective learning strategies and were more likely to self-regulate the time spent each week on the course, resulting in successful course performance (N=256). These students were also more confident in the technology used in the online environment, and consequently were more satisfied with the course overall than those who had not previously taken an online course.

Cole, Shelley, and Swartz (2014) conducted a three-year study of student satisfaction in both undergraduate and graduate business courses. With an impressive 92% response rate, respondents were characterized by gender, age, and course classification, undergraduate or graduate (N=553). Those who participated in the short survey were offered extra credit for their contribution. The mixed methods research asked students to first rate their satisfaction with their online course, using a five-point Likert scale, where 0 represented very satisfied with the course and 4 very dissatisfied. Second, students were then asked to articulate those factors that contributed to their satisfaction/dissatisfaction in the online course they were taking. For answers to the first question, researchers used independent t-tests to evaluate the data. The mean scores for male and female were 1.4 and 1.3, respectively; the mean scores for those younger than 35 and older than 35 were 1.4 and 1.24 respectively; the mean for undergraduate and graduate were 1.19 and 1.23, respectively. The second question of the survey revealed six major factors that contributed to a student’s satisfaction level in the online environment. Students most satisfied typically cited convenience, course structure, and learning style as reasons for preferring
the online mode. Those dissatisfied predominantly mentioned lack of interaction with the instructor and lack of interaction with other students. Additionally, some students also stated online course structure and the online learning platform as the primary reasons for their dissatisfaction. This study brings to light the importance of satisfaction characteristics to consider in the online environment. Namely, convenience, structure, learning platform, interaction, and learning style have been identified as important attributes to consider in a course conducted using a remote environment (Cole et al., 2014).

Armstrong (2011) conducted a study exploring students’ perceptions and experiences in online classes. Specifically, the study focused on what student characteristics helped motivate them to succeed in an online course, including identifying positive and negative aspects of online courses, what students’ initial perceptions of online classes were in terms of difficulty as compared to traditional classes, and what students believed professors could do to improve the online course experience. The researcher’s findings were categorized into five themes: (a) communication from the professor was of greatest importance when shaping a students’ perception of online courses, (b) how the technology was implemented by the instructor mattered more to students than what technology tools were used, (c) organized course structure was directly related to a students’ ability to succeed, (d) the course structure and type of assessments influenced the students’ approach to learning and their success in the online course, and (e) students failed to utilize the available library databases when completing assignments, opting for nonacademic sources instead (Armstrong, 2011). Armstrong’s (2011) findings complement those of Cole et al. (2014) in that interactions between the student and the instructor, course structure, and learning platform all contributed to student satisfaction in a web-based platform.
In another research study evaluating student satisfaction in an online course, Lee’s (2014) objective was to help improve student learning in the online environment (N=81). The instrument used was designed by the researcher to measure student satisfaction by considering both human and design factors, where human factors were questions related to the professor and the graduate assistant, and design factors were questions related to the course structure and the use of technology. Researchers found that human factors impacting student satisfaction were mostly related to the professor’s mastery of the course material as well as familiarity with current trends in the field. Also, students wanted confidence in fair grading of their assignments as demonstrated by constructive feedback from the graduate assistant grading the material. From both the professor and graduate assistant, students were most satisfied when they received prompt replies to their questions. In a study conducted by Lee, et al. (2011), researchers had similar findings when evaluating student perceptions (N=110). Additionally, they found student satisfaction was directly related to the amount of support they received while taking an online course, including continual feedback from the professor throughout the course (Lee et al., 2011).

For the design factors, Lee (2014) found the most significant element of a student’s satisfaction in the course was from clearly defined course expectations, along with detailed guidelines and rubrics. Also, easy to use technology was important, especially since not all of the participants in the study described themselves as being adept with technology. In regard to proficiency, Lee et al. (2014) also found that beginning the course with an online orientation to help students navigate the online course improved student satisfaction.

Kuo, Walker, Belland, and Schroder (2013) conducted a study to determine predictors for student satisfaction in 11 online classes (N=111). Specifically, their study considered how a student’s satisfaction in an online course was impacted by their interactions – Moore’s (1999)
student-to-student, student-to-instructor, and student-to-content interactions – as well as by their level of internet self-efficacy, and their self-regulated learning practices. The data collected was from three questionnaires. Similar to studies that have also considered student performance, the researchers collected demographic data related to age, gender, course level, and self-reported hours of time spent each week on the online course (Driscoll et al., 2012, Xu & Jaggars, 2014). The second and third questionnaires both used a seven-point Likert scale to determine the level of internet self-efficacy and self-regulating learning, respectively. For internet self-efficacy, a scale developed by Eastin and LaRose (2000) was used for the study. For self-regulated learning, Pintrich et al.’s (1993) Motivated Strategies for Learning Questionnaire was used, which is based off of Zimmerman and Martinez-Pons (1986) self-regulated learning strategies. Using multiple regression, results from the study confirmed prior research in that the learner-content interaction is the strongest predictor for student satisfaction in an online course (Armstrong, 2011; Cole et al., 2014). Similar to Driscoll et al. (2012), this research also confirmed that internet self-efficacy significantly contributed to student satisfaction in an online delivery.

Summers et al. (2005) found that nursing student satisfaction in a statistics course significantly differed by environment in that online students, overall, were more dissatisfied with the course than their counterparts in the face-to-face offering, even though their instructor, exams, and, ultimately, performance was the same (N=38). The researchers suggested that the dissatisfaction was the result of the missing learner-instructor contact. Students were unable to directly interact with the instructor, resulting in the perception that they lacked sufficient understanding of the content. These researchers suggested design strategies for online courses to minimize the distance between student to instructor and student to content through “(1) course
planning and organization, (2) verbal and nonverbal presentation skills, (3) collaborative teamwork, (4) questioning strategies, (5) subject matter expertise, and (6) involving students and coordinating their activities at field sites” (Summers et al., 2005, p. 239)

As in the case of Summers et al. (2005), Carter and Emerson (2012) also found that student satisfaction of those enrolled in an online Principles of Economics course was significantly lower than those enrolled in a similar face-to-face course (N=176). Specifically, the traditional course included experiments that were hands-on and in-class while those in the online version were fully computerized. Using a five-point Likert scale, students were asked to rank both the level to which the experiments helped them to learn the subject material as well as their likelihood in taking another economics course. Although students performed similarly in both delivery modes, students in the online sections were more dissatisfied with the course, specifically disregarding the computerized experiments as a means for learning the material. Not surprising, this control group was also less likely to want to take another economics course.

Simpson and Benson’s study (2013) used Aman’s Satisfaction Questionnaire to assess student satisfaction in online courses (N=157). The 30 question survey, using a five-point Likert scale, considered standards related to learning objectives, assessment and measurement, instructional materials, learner interaction and engagement, and course technology (Simpson & Benson, 2013). Using a one-way ANOVA, student satisfaction in the online courses was found to be significantly higher in three areas. First, older students, those 25 to 35 years old and those 55 years or older, were more satisfied with the online environment rather than the traditional face-to-face environment. Second, students who described themselves as more comfortable with distance learning were also more satisfied with the online delivery mode over face-to-face. Third, students who had completed ten or more previous online courses were also likely to prefer
the online setting over the traditional. As Carter and Emerson (2012) suggested, Simpson and Benson (2013) concurred that best practices should be implemented when using technology in online courses to enhance student perception and overall satisfaction.

Bergstrand and Savage (2013) considered student evaluations in terms of student satisfaction in both the online and traditional classrooms. In an introductory sociology course, the findings of this study suggested that students in online classes were overall less satisfied with the class. Specifically, students enrolled in the online environment found they learned less, believing the instructors to be less effective and less respectful than those who taught the same course face-to-face (Begstrand & Savage, 2013). It is of wonder, then, why the explosion of online course offerings continues? Bergstrand and Savage (2013) did note that while this was the summary of their findings, it was not a consistent assessment across all online classes. Some faculty members have better designed online courses than others and therefore have higher overall student satisfaction.

Simpson and Benson (2013) suggested that faculty peer review of online courses could boost student satisfaction, and while this is sensible, their research did not show significance between higher student satisfaction rates in online courses that had undergone these kinds of reviews and those that had not. As in the study by Summers et al. (2005), students in the online courses of the Bergstrand and Savage’s (2013) study evaluated their class more negatively than those in the equivalent face-to-face version of the course, even though their perceptions did not necessarily reflect their overall performance in the class.

Using students enrolled in both an online and traditional face-to-face upper level histology science course (N=44), Schoenfeld-Tacher, McConnell, and Graham (2001) analyzed the impact of Moore’s (1999) theory of distant interactions on student learning through a mixed-
methods study. The baseline for students’ incoming knowledge on the subject material was ascertained by a pre-test at the start of the course. The qualitative component was collected throughout the semester from student and instructor questions observed and recorded in three possible environments: classroom lecture, online discussion board, and student-organized online study sessions. Each time any of these three environments occurred, questions asked during the particular session were categorized as content, administrative, management, or social. Results from the coded questions were analyzed using t-tests, ANCOVA and ANOVA. Overall, in both environments, findings showed that content-initiated questions by both instructor and students usually occurred during classroom lecture rather than the online discussion boards. Furthermore, regardless of environment, students were more likely to initiate the questions first, not the instructor. The online students, however, asked more questions than those in the face-to-face class, but this was expected given the remote environment. For the quantitative component, test scores were compared for each delivery mode. The pre-test showed no initial difference in prior knowledge of course content for either environment. However, final results showed students in the online course learned significantly more than those in the traditional face-to-face course, based on the multiple choice test given to all forty-four students in the study. As with so much of the literature regarding online content, this study adds to the debate of which mode is better by suggesting that online is actually better than traditional.

Understanding what elements contribute to student satisfaction, specifically in online courses, can help to improve the design of classes in the online environment and enhance a students’ overall experience, strengthening retention and performance rates. Moreover, better identification of pedagogical practices can serve to diminish the negative impact of transactional distance on a student’s ability to succeed in the online delivery mode.
Summary

This literature review has provided an examination of the current debate in the quality of online education. The mixed results of student course outcomes in the online environment relative to those in the traditional face-to-face suggest that more research is still necessary to better ascertain differences in student achievement, if any, in the online environment over the traditional face-to-face environment. As Moore’s transactional theory suggests, distance is an important factor that must be considered and yet the literature still has no definitive answer into the part it plays in the student learning environment. Moreover, studies based on Pintrich’s (2004) SRL model suggest that student achievement directly relates to student satisfaction in that students who are more satisfied with a course are also more likely to be committed to the necessary self-regulated learning practices that must be employed to succeed, even if the course environment is remote (Artino & McCoach, 2008; Bolliger & Martindale, 2004; Reinhart & Schneider, 2001).

Although this review has focused on four-year institution studies, many of which had low sample sizes, the most compelling studies have been conducted at the community college level (Hachey et al., 2012; Hachey et al., 2014; Jaggars et al., 2013; Wladis et al., 2014; Xu & Jaggars, 2014). It is evident that more studies like those at the community college should be modeled at universities to better understand the factors related to student achievement, and my research has contributed to this gap in literature. Moreover, one strikingly noticeable gap is the lack of STEM studies. Literature addressing online STEM is severely lacking. Typically, only the soft sciences have been assessed in terms of delivery mode, either at the community college level or the university. The research study I conducted investigating a geology survey course has contributed to this limitation in literature.
Although more recent research suggests that today’s students have adapted to the online class in terms of performance, their overall performance and satisfaction have not (Bergstrand & Savage, 2013; Carter & Emerson, 2012; Simpson & Benson, 2013; Summers et al., 2005). It is important for the strategic planning of higher education institutions to consider pedagogical design of online courses in all fields of study, including STEM disciplines. Moreover, better identification of instructional practices can serve to diminish the negative impact of any transactional distance on a student’s ability to succeed in the online classroom. Indeed, it is essential for post-secondary institutions and their mission to educate society to assess how content is delivered in a technological environment in terms of the student and their interaction with the instructor, as well as their achievement and satisfaction with the curriculum in their field of study.
CHAPTER III

METHODS

Through the lens of the SRL model and Moore’s theory of transactional distance, this research considered an online physical geology course in terms of learning outcomes and student satisfaction, especially pertaining to the interactions between student, instructor, content, and technology. Although studies have been conducted to compare the two environments, research results are not definitive as to how physical distance, specifically the online platform in this study, impacts student achievement, nor how it affects perceptions regarding student satisfaction. Because the design of this study was to determine both factors that predict student achievement in an introductory survey geology course by type of classroom environment, online versus traditional face-to-face, and perceptions that influence student satisfaction in the online delivery mode, a concurrent triangulation mixed methods approach was used as the design of this study (Creswell, Clark, Gutmann, & Hanson, 2003).

Research Design

According to Creswell (2014), the definition of a general mixed methods study is one that employs both a quantitative and a qualitative approach. Analyses are rendered using both closed-ended and open-ended data (Creswell, 2014). Campbell and Fiske (1959) first introduced multitrait-multimethod (i.e., mixed method) research, arguing that findings were more convincing from data collected through multiple methods. Their reasoning was that data analyzed using both quantitative and qualitative approaches would yield the same conclusion, even though the research design for collecting data from each was fundamentally different.
Specifically in this study, a concurrent triangulation mixed methods strategy was used by collecting and examining the quantitative and qualitative data separately, and integrating the findings during interpretation – the qualitative data was used to broaden the perspective of the quantitative data results (Creswell et al., 2003).

In this study, the quantitative method used archival student achievement data, while the qualitative method explored characteristics leading to student satisfaction. For the quantitative component, this study sought to determine whether those students in an online environment performed differently than those in a traditional environment by measuring achievement through two proctored exams and the final grade, while controlling for incoming grade point average (GPA). The dependent variable was each of the tests and the final grade, while the independent variable was course environment – online or traditional face-to-face – using GPA as a covariate. Additionally, student characteristics were collected to determine whether certain descriptors could be used as predictors for success in an online environment, namely, GPA, prior success in online courses, type of declared degree program (STEM and non-STEM, Geology and non-Geology), student classification, and semester course load. For this analysis, the dependent variable was the final grade in the online course, and the independent variables were the various descriptors. Details of the statistical analyses are further described in the Data Analysis section.

Because I knew the instructors of record for both the online and traditional face-to-face courses, I asked if they would be willing to share their data with me for this study. They both agreed, and after my dissertation committee and IRB approval (Protocol #16x-201, exempt under 45 CFR 46.101(b)(#2, 4)), this study was possible because of the data they were willing to provide. Even though I knew the instructors, if owners of the data are willing to release their data and the data remains FERPA compliant, such a study could be replicated.
For the qualitative component, a focus group was conducted to ascertain perceptions of satisfaction in the online course, including those descriptions of motivational strategies that were implemented by students, if any, to maximize learning outcomes, as well as what influence the learning platform had on a student’s level of achievement. All students enrolled in the course were emailed, asking for volunteers to participate in the interview. The email specifically detailed that whether or not they choose to participate in the focus group would have no bearing on their grade in the course. Those respondents who agreed to participate were given nominal compensation in the form of a gift card for their time.

The results of this study may be generalizable to other post-secondary students taking a similar online survey geology course on natural resources, waste disposal, and climate change in predicting performance outcomes and characteristics determining satisfaction levels. In addition, this study described those perceptions that influenced student satisfaction in an online survey geology course related to the research questions of this study.

**Qualitative Research Question**

How do students describe their level of satisfaction in online classes?

1. What perceptions contribute to student satisfaction in online classes?

2. How do students describe their self-regulating learning practices that contribute to student success in online classes?

3. What do students believe should be implemented to improve the delivery of online classes?

4. How do students’ perceptions and expectations of online classes differ from traditional, face-to-face, classes?

**Quantitative Research Question**
Does student performance differ by classroom environment: online versus traditional face-to-face?

**Hypotheses**

Hypothesis One: There is no significant difference in the mean score of test one by type of classroom environment (online or traditional) when controlling for incoming GPA.

Hypothesis Two: There is no significant difference in the mean score of test two by type of classroom environment (online or traditional) when controlling for incoming GPA.

Hypothesis Three: There is no significant difference in the mean score of the final grade by type of classroom environment (online or traditional) when controlling for incoming GPA.

Hypothesis Four: There is no significant relationship in the mean score of the final grade and a group of predictor variables, including GPA, prior online experience, major as categorized as STEM and non-STEM, major as categorized as Geology and non-Geology, number of completed student credit hours, and current course load (number of credit hours for the semester).

**Population, Sample, and Participants**

The sample for this study came from students at a midsize public Southern university enrolled in Environmental Geology - Resources (Geol 105), either in an online or traditional environment. Students at any university, not just this particular university, who have chosen to take a survey science course such as Geol 105, do so to partially fulfill the general science requirement for their degree programs. Students are typically from both science, technology, engineering, and math (STEM) majors and non-STEM majors. An additional unique quality of Geol 105 is that, of the four survey Geology courses offered at this university, it is the only one that is also a core course required for those students who have declared Geology as their major.
Although the online version of this course is typically taught in the fall semesters and the face-to-face in the spring semesters, both used the same textbook and the same test banks. The course was initially offered online in the Fall semester of 2011, while the face-to-face was first offered in the Spring of 2010. Because the initial Spring 2010 semester used a different textbook, it is not included in this study. Rather, the first semester of the traditional course offering used for this study was Spring 2011. While the instructors for each platform were different, both had several years of experience teaching at this university in both the online and traditional face-to-face environments for the Geology department, and both had similar teaching evaluations, including for Geol 105. Moreover, the online offering of Geol 105 has been taught by the same instructor for the four fall semesters that were used in this study, and, likewise, the traditional offering was taught by the same instructor for the five spring semesters that were used in this study. Further, the PowerPoint lectures, used as the primary mode of instruction, was the same for both the online and the traditional, face-to-face, offerings of Geol 105, other than the online also incorporated audio to the lectures to compensate for the lack of face-to-face interaction.

The face-to-face version of this course enrolled approximately 200 students in the spring semesters. The student population included ethnically diverse individuals of all classifications – freshman, sophomore, junior, or senior – and was primarily comprised of the traditional aged college student, 18-22 (http://nces.ed.gov/programs/coe/indicator_csb.asp). For this study, data was collected for the Spring semesters of 2011, 2012, 2013, 2014, and 2015 for a total of approximately 1,000 students. The teaching method used was two and a half hours of lectures per week for fourteen weeks in the regular semester, either in three fifty-minute or two one-hour-and-fifteen-minute sessions. The course was taught using traditional methods, including
classroom lecture, PowerPoint, and short videos. Although attendance was taken and part of the students’ overall grade in the course, no homework was given in the class, only reading assignments from the textbook. Though not required, students also had the option of simultaneously enrolling in the lab that is associated with the course. Three paper tests were given during the semester, each of which was comprised of only multiple choice questions.

The online Geol 105 course was limited to approximately forty-five students per regular fourteen week semesters. Like the face-to-face course offering, the student population included ethnically diverse individuals of all classifications – freshman, sophomore, junior, or senior – and was primarily comprised of the traditional aged college student, 18-22 (http://nces.ed.gov/programs/coe/indicator_csb.asp). For this study, data was collected for the Fall semesters of 2011, 2012, 2015, and 2016, for a total of close to 200 online students. The Fall semester of 2013 was be omitted from the data collected because a different instructor was responsible for the online section of only this semester.

Because the two populations for this study were significantly different in size, the traditional face-to-face course had more than triple the students enrolled than the online course, it was preferred to determine the same, optimal, sample size that should be used for each population, minimizing the Type I and Type II errors that are characteristic of hypothesis testing (Hinkle et al., 2003). That is, the researcher must consider when statistical analysis might reject a true null hypothesis (Type I error), or fail to reject a false null hypothesis (Type II error). A power test provides a confidence level for rejecting a false null hypothesis (Hinkle et al., 2003). The quantitative approach of the study used G*Power (Faul & Lang, 2007), a software tool developed to calculate an optimal balanced sample size in statistics based on power, such that the number in each of the two populations being compared is equal. For ANCOVA, using, a type I
statistical error ($\alpha$) of 0.05, a power ($\beta - 1$) of 0.95, and a medium effect size ($\omega^2$) of 0.25, G*Power determined that an optimal minimal sample size for this study was 210 (N=210). Therefore, hypotheses one, two and three required using a minimum of 105 students from each learning platform – online and traditional face-to-face, for a total of at least N=210 for this study.

For OLS, using, a type I statistical error ($\alpha$) of 0.05, a power ($\beta - 1$) of 0.95, a medium effect size ($\omega^2$) of 0.15, and six predictors, G*Power determined that an optimal minimal sample size for this study was 146 (N=146). Therefore, hypothesis four will use at least 146 from those enrolled in the online Geol 105 course. Details on each of the statistical methods are discussed in chapters four and five.

The online course was administered through the Blackboard online course management system (http://www.blackboard.com/). The teaching method for the course was comprised of weekly units using exclusively PowerPoint, similar to those used in the traditional face-to-face course, but with audio added to each file. Students were expected to review the PowerPoint slides as well as read the associated textbook material. To ensure students regularly logged onto Blackboard and read the current week’s material, open-book quizzes were given for each unit to monitor students’ progress. Students were permitted to take the quiz as many times as they wanted, which drew from a pool of questions, and only the last quiz submission for that week’s quiz grade was recorded. Like those in the traditional classroom, students enrolled in the online version of this course could also enroll in the accompanying lab. Two online exams were given, also comprised of only multiple choice questions. The exams were taken on Blackboard and proctored by the university’s outreach department.

The course was fully accredited by the Southern Association of Colleges and Schools (http://www.sacscoc.org/principles.asp) in fulfillment of the requirements for college credit in a
survey course covering resources in Environmental Geology: natural resources, waste disposal, and climate change. The material covered in both courses at this university would be the same as what would be found in any resources environmental geology course offered at the college level.

**Instruments**

For the quantitative approach of this study, the measurement instruments were the two course exams given during the semester and the final score for the online course, and the three course exams given during the semester and the final score for the traditional, face-to-face, course. Because tests 1 and 2 of the traditional course covered the same material as test 1 in the online course, the average of test 1 and 2 in the traditional course was used to compare against test 1 in the online course. Therefore any reference to test 1 in this study pertains to test 1 in the online course and the average of tests 1 and 2 in the traditional course. Likewise, any reference to test 2 in this study pertains to test 2 in the online course and test 3 in the traditional course.

The assessments for both modes of instruction were exclusively multiple choice questions developed by each of the instructors to test a student’s knowledge of the geologic principles covered throughout the course. When the online course was initially in the development stages, the instructor for the face-to-face course provided all the necessary material to the online instructor, including PowerPoint slides and test bank questions. The online instructor used this as a baseline for developing the material for the online version of the same course. For the lectures, audio was added to the PowerPoint lectures. For the exams, pools of questions for both the quizzes and tests were created. These questions are similar to that of the face-to-face tests, other than they are tailored to the wording of the online instructor. Although questions for the exams are not necessarily identical, the questions from both environments could be aggregated by topic. Constructing a test in this way would be similar to the same instructor
writing comparable, but different, exams from semester to semester in which the course is taught.

**Data Collection**

For the qualitative approach of this study, a small subset from only those students enrolled in the online Environmental Geology – Resources class for the Spring 2016 semester participated in the focus group interview. During the Spring 2016 semester, the instructor of record sent out an email to students describing this research study to solicit volunteers to participate (see Appendix A). The email stated that participation in no way affected their grade in the course, but those who did choose to participate would be given nominal compensation in the form of a gift card for their time. Of those students who responded, the intent of this research study was to extend an invitation to seven to ten students, since seven to ten has been identified as the ideal number of participants in a focus group (Krueger & Casey, 2000). Students were to be chosen such that there was fair representation of each grade range, from superior to failing. Students who responded first to the invitation to participate would be considered first, with secondary consideration given to how they were currently performing in Geol 105. As discussed in the results section, only four students responded and so the focus group only had four participants.

The focus group is designed as a semi-structured group session where questions are asked by the researcher, and students are encouraged to be honest in their answers regarding their experiences in the online class. Questions that were asked by the researcher are included in Appendix B. Because of the semi-structured nature of the interview, the students in the group session were also encouraged to engage in a general discussion of their practices and opinions regarding the online environment. Data was collected through multiple sources, including the
interview itself, the audio-taping of the interview, and notes during the interview. Triangulation of the data findings was accomplished using three individuals’ independent coding of the focus group transcription.

**Procedure**

Approval for this study was first obtained by the researcher’s dissertation committee members, followed by approval from the university’s Institutional Review Board (Protocol #16x-201, exempt under 45 CFR 46.101(b)(#2, 4)).

**Quantitative Procedure.** Because the quantitative part of the research study used archival data, the two Geol 105 instructors were contacted and agreed to provide their gradebooks for the semesters used in the study. They also agreed to provide the university student identification numbers, along with students’ corresponding scores on the two tests as well as the final score and letter grade in the course. Finally, the director of the university’s Institutional Research, Effectiveness and Planning (IREP) department was contacted and agreed to provide various characteristics for each of the students who have taken Geol 105. At this university, the IREP department is responsible for collecting and archiving data related to student characteristics at the beginning and end of each semester, providing a snapshot of a description of the student enrollment. For the predictor variables to be meaningful for this study’s statistical analyses, it was important that the data collected was for the incoming fall semesters in which students had taken their online course: 2011, 2012, 2014, or 2015. Using the university student identification numbers, IREP was able to provide the following information of each student for the incoming semester in which they had taken Geol 105: GPA, the number of previous successful online courses, declared degree program, semester course load, completed student credit hours.
The Family Educational Rights and Privacy Act of 1974. The Family Educational Rights and Privacy Act of 1974 (FERPA) does not allow student educational records to be released to a third party, nor does it allow for records to contain identifiers ascertaining a student’s identity (http://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html). Because the gradebooks, as provided by the instructors with student identification numbers, were protected under FERPA, electronic files containing the Geol 105 gradebooks for the nine semesters used in this study were housed using a FERPA-compliant university cloud service known as Box. The chair of the Geology and Geological Engineering department created two private Box folders for the data collection, one for the online Geol 105 files, and one for the traditional Geol 105 files. As owner of these folders, the chair first granted each instructor access to the folder corresponding to their class environment, online or traditional, and the IREP director was granted access to both. Each instructor then uploaded their files to their corresponding Box folders, followed by the researcher’s request to IREP to add the necessary data for the research study. The IREP director then appended the student characteristic data to each of the students in the gradebook files on Box. Additionally, the IREP director grayed out the student identification numbers such that any identifiers to students, namely the student identification number, was no longer listed in the files. Finally, the IREP director created a separate Box folder with the uploaded FERPA protected files and granted access to the researcher so that the quantitative phase of the research could be conducted.

Qualitative Procedure. The qualitative approach of the study was to assemble the focus group from those enrolled in the Spring 2016 online Geol 105 course. Students were asked during the term to respond whether they would consider participating in a small focus group at the end of the semester, which would require them to answer questions in a semi-structured
interview environment. Additionally, students were informed that the questions and any discussion topics were only related to those factors that affected their level of satisfaction and perceived performance in the class, and did not have any impact on their overall grade in the class. Students were also offered an incentive in that those who participated would be given a gift card for their voluntary participation in the study (Vaughn, Schumm, & Sinagub, 1996). The email interchanges with students are listed in Appendix A.

Initially, the thought was that from the pool of students who agreed to participate in a focus group, seven to ten would be purposefully selected based on their midterm grade. All grade ranges from outstanding to failing performance ideally would be represented in the focus group. Students were told upfront that participation and responses would be confidential and would not have any bearing on their final grade in the course. Although the exact number of participants is unnecessary in designing a credible focus group, characteristics of those student recruited is important for trustworthy findings, and my study meets this requirement (Vaughn et al., 1996).

The group session was approximately one hour, and was conducted as a semi-structured interview. Questions for the interview are listed in Appendix B. The researcher audio-recorded the interview as well as took notes on the group discussion in order to strengthen reliability and internal validity of the data.

Implementing focus group interviews typically follow a protocol with particular components (Vaughn et al., 1996). The focus group conducted for my study implemented the following procedure:

1. Identified objectives of the interview.
2. Selected a moderator who understood the objectives.
3. Chose respondents to participate in the group.

4. Identified a location for the interview.

5. Outlined questions to be asked in the interview.

6. Conducted the interview.
   
   a. An explanation of the reason for the focus group was given.
      
      i. The intent of the focus group was to explore their perceptions of the online environment, specifically in how satisfied they were taking a course through the web as opposed to the traditional classroom.
      
      ii. Periodic questions were asked throughout the time, but the focus was on what they chose to discuss as it related to their experiences in Geol 105.
      
      iii. Everyone was expected to engage in discussion, but not everyone was required to answer all of the questions.
   
   b. Students were asked to introduce themselves so other students could refer to them by name.
   
   c. The focus group began by my asking the first question in Appendix A, followed by student discussion of this question and other issues related to their experiences in the online class.
      
      i. As discussion continued, I took notes. Periodically, students were asked to embellish on their statements in order to better understand their perceptions.
      
      ii. When the discussion stalled, I asked another question listed in Appendix A.
iii. When the discussion would get off topic, I brought it back to the issue of satisfaction in the online environment by asking another question listed in Appendix A.

iv. If one student was not participating in the focus group, I specifically asked this individual a question, either one that related to the current discussion or one that was listed in Appendix A.

7. Interpretation of the data.

**Data Analysis**

**Statistical Test.** In the research conducted by Driscoll et al. (2012), several control variables were introduced in their study: self-reported GPA, current number of credit hours, weekly number of hours worked in a paying job, year in school, and number of previously taken online courses. Because GPA has been shown to be a significant covariate (Driscoll et al., 2012; Frantzen, 2012; Hachey et al., 2014), ANCOVA, rather than ANOVA, was the preferred statistical method for this research study to test for significant relationships in the data regarding hypothesis one, hypothesis two, and hypothesis three.

The analysis of covariance (ANCOVA) is a statistical tool that combines the power of regression and analysis of variance (ANOVA), while controlling for unwanted effects from dominant variables (Hinkle et al., 2003). The ANOVA statistical procedure is an optimal choice when comparing just two sample means from a population, such as for this study comparing the means of student achievement by two classroom environments. ANOVA, however, does not consider any control variables that may be influencing the calculated means. In such a case, ANCOVA is preferred because the covariate, or the extraneous variable effect, is separated out to allow for a more accurate determination of the primary independent variable by statistically
controlling how much this highly correlated covariate factor influences variations on the measurement (Hinkle et al., 2003). In essence, this variable by itself statistically contributes to the variance; therefore, it is controlled such that it is included but does not dominate the calculation.

For hypotheses one and two, the dependent variable was the test score (test one for hypothesis one and test two for hypothesis two), and the independent variable was the classroom environment (online or traditional modes of delivery). For hypothesis three, the dependent variable was the final score, and the independent variable was the classroom environment (online or traditional). Since a students’ natural aptitude may influence their performance on a test as shown in past research (Driscoll et al., 2012; Frantzen, 2013; Hachey et al., 2014), incoming GPA was used as the covariate to help determine the true statistical effect of the test scores by the classroom environment. ANCOVA assumes linearity between test scores and GPA, and non-linearity between each of the independent variables and GPA.

Additionally, in order to meet the assumption of homogeneity of variance, which was a potential issue with an unbalanced sample size (approximately 200 from the online students and approximately 1,000 from the traditional face-to-face students), a technique known as propensity score matching (PSM) was initially used to reduce the sample size of the face-to-face environment to 171, the same as the sample size of the online Geol 105 course. The total sample size (N=342) also met the minimum sample size of 210 as determined by the power test (Adelson, 2013). PSM is a technique that uses propensity scores to partition the data into groups, creating subsets that have been statistically aggregated by a matching set of conditions. PSM is preferred so that the means calculated between the two groups are equally weighted rather than biased by the traditional classroom’s larger sample size. By trimming the data, the
determined outcome is truly reflective of the mean between the two groups: online and face-to-face. The MatchIt package offered in the open-source programming language R was used for PSM (Randolph, Falbe, Manuel, & Balloun, 2014). Initially, all other statistical analyses used SPSS 22 (http://www-01.ibm.com/software/analytics/spss/). Because of assumption violations that will be discussed later, some statistical analyses used R for the statistical computing instead (https://www.r-project.org/).

According to Hachey et al. (2014), however, GPA alone is insufficient in terms of covariate considerations for the online environment. Their study showed that GPA was a predictor for success, but only for students who were taking an online class for the first time. The predictor for students who had previously taken online courses, however, was solely their success in those previous courses, regardless of their incoming GPA. Thus, for hypothesis four, an ordinary least squares (OLS) linear regression model, often referred to as multiple regression when more than one independent variable exists, was the most appropriate choice of analysis. The sample size for the OLS was 171, which also met the minimum required sample size of 146 determined by the power test.

Multiple regression is a univariate analysis tool that allows for many variables to be used as predictors for a particular criterion. Specifically, the level of prediction of one dependent variable (Y) can be determined by introducing several independent predictor variables (X₁, X₂, …, Xₖ). Using calculated regression coefficients (b₁, b₂, …, bₖ) for the slopes, and a regression constant (a) for the y-intercept, all are combined into what is known as a multiple regression equation (Hinkle et al., 2003):

\[ \hat{Y} = b_1X_1 + b_2X_2 + \ldots + b_kX_k + a \]

\( \hat{Y} \) = predicted single criterion, or single dependent variable
\(X_i\) = predictor variable, or one of the dependent variables

\(b_i\) = regression coefficient, or slope of the line for its paired predictor variable

\(a\) = regression constant, or y-intercept

Each predictor variable’s regression coefficient \((b_i)\) in the multiple regression equation is calculated to represent the corresponding predictor variable’s proportion of variance, which effects variance in the independent variable. Furthermore, for the OLS formula to accurately predict an outcome, it is important to verify that each predictor variable has low correlation with the other predictor variables. If two predictor variables do correlate, also known as multicollinearity, then the variance of each would explain the same variance in the independent variable. This association is not allowed since variance can only be quantified once (Hinkle et al., 2003).

Because studies have shown that other predictors exist besides GPA shown in Driscoll et al.’s (2012) study (Hachey et al., 2014; Wladis et al., 2014), hypothesis four used an OLS that considered not only GPA but other possible predictor variables for success in the online environment as well. For hypothesis four, the dependent variable was the final grade in the online class, and the independent variables were GPA, the number of successful previous online courses, student classification – Freshman, Sophomore, Junior, Senior – as determined by completed student credit hours, total number of student credit hours taken in the same semester as the online Geol 105 course, and the student declared degree program – categorized as STEM or non-STEM major and Geology or non-Geology major. All statistical analyses for hypothesis four used SPSS 22 (http://www-01.ibm.com/software/analytics/spss/).

**Focus Groups.** A focus group is a discussion session conducted by one person to a group of people who have been brought together to deliberate on a topic that they all have in
common (Gall, Gall, & Borg, 2007). The optimal size of a focus group has been identified as seven to ten, where the atmosphere is meant to be relaxed and to explore issues related to the topic of discussion (Krueger & Casey, 2000). Typically, the interview is a semi-structured design, where the interviewer has some initial questions to ask the group in order to facilitate conversation between the interviewer and the individual, as well as between each of the individuals participating in the session. This set up for a focus group interview allows for more stimulated discussion than a one-on-one semi-structured interview, and it enables the interviewer to serve not only as a facilitator, but also in a more non-directive role as an observer (Gall et al., 2007).

For the qualitative component of this study, I conducted a focus group from students enrolled in the Spring 2016 semester Geol 105 online class, which was held on the main campus of the university to explore the research questions and to identify those themes that emerged from the discussion in terms of student satisfaction. The purpose for adding the focus group element was to broaden the understanding of the results of the quantitative analysis, offering a thicker description of student performance through the lens of student satisfaction. As the moderator for this interview, I have both taught online courses as the instructor of record and taken online courses as a graduate student. As a result of my own experiences, I have long been interested in those issues that influence student satisfaction in the online classroom, which, in turn, have influenced student performance in such an environment. Because online classes are remote, instructors typically only have opportunity to gauge student performance from assignments given in the course, with little to no knowledge of how satisfied their students are, particularly how student satisfaction levels have influenced achievement. Conversely, students in online courses are also limited in the level and expediency of communication they have with
their instructor and other students. When a student has an issue or question, for example, they cannot expect immediate feedback as they would in a face-to-face classroom. As a knowledgeable moderator in the online environment, my own experiences helped in facilitating discussion with the students.

Table 1

*Focus Group Semi-Structured Interview Questions*

<table>
<thead>
<tr>
<th>Thematic Category</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convenience</td>
<td>1. Why did you enroll in the online version of Geol 105 rather than the traditional, face-to-face version?</td>
</tr>
<tr>
<td>Structure</td>
<td>2. How many online classes have you taken before?</td>
</tr>
<tr>
<td></td>
<td>3. How does Geol 105 compare to these other courses?</td>
</tr>
<tr>
<td>Learning Style</td>
<td>4. When you reflect on your learning style, do you think online was a good choice, or would the traditional face-to-face have been better? Why?</td>
</tr>
<tr>
<td>Interaction</td>
<td>5. Did you find your instructor provided prompt and helpful responses to your questions?</td>
</tr>
<tr>
<td>Platform</td>
<td>6. How comfortable were you using Blackboard?</td>
</tr>
<tr>
<td></td>
<td>7. Was any other technology, such as YouTube or another website, used besides Blackboard?</td>
</tr>
<tr>
<td>All</td>
<td>8. What would you identify as the positive characteristics of having taken this course online?</td>
</tr>
<tr>
<td></td>
<td>9. What would you identify as the negative characteristics of having taken this course online?</td>
</tr>
<tr>
<td></td>
<td>10. Would you take another online class? Why or why not?</td>
</tr>
<tr>
<td></td>
<td>11. Overall were you satisfied with Geol 105?</td>
</tr>
<tr>
<td></td>
<td>12. What would have made Geol 105 even better?</td>
</tr>
</tbody>
</table>

It is important that the focus group discussion does not stray from the topic (Patton, 2002). Table 1 delineates the questions that will be asked to keep the discussion on the topic of student satisfaction. In the study conducted by Cole et al. (2014) on factors influencing student satisfaction in online courses, the researchers identified six themes that emerged from the data: convenience, interaction, structure, learning style, and platform. In light of these findings, this
The study’s focus group questions in the semi-structured interview were purposely devised around these classification schemes, also noted in Table 1.

The focus group was transcribed and the data analyzed for those emerging themes that helped to answer the research questions, including those factors that students identified as ways to enhance their overall experience and satisfaction in an online class, both in how the class was deployed as well as changes they could have made in their own behaviors to improve their self-regulated learning practices.

To control for unwanted bias I might bring to coding the data, I first used computer-assisted qualitative data analysis software (CAQDAS) to organize and manage the transcription before analyzing for recurring regularities in the data. CAQDAS is a tool used in qualitative studies to efficiently code, compare and link data, which often is a tedious task for the researcher (Patton, 2003). In their study comparing qualitative data coded by hand and coded by CAQDAS, Rademaker, Grade, and Curda (2012) found that the visual representation of coded CAQDAS data helped them better understand the occurring themes in their data, this despite the different and unique background of each of the researchers. For my study, NVivo 11 was used (http://www.qsrinternational.com/), a low cost CAQDAS for students. Once the qualitative data was entered into the CAQDAS, I initially looked for those themes identified by Cole et al. (2014). I also considered additional categories from the CAQDAS results that had internal homogeneity and external heterogeneity. Patton (2002) defines internal homogeneity as data that relates in a meaningful way and external heterogeneity as data that has distinct difference. From these results, I identified those themes with substantive significance that emerged from the focus group relating to student satisfaction in the online Geol 105 class. Finally, I had two additional coders, one who is a specialist in online instructional design and a second who has coded other
focus groups, also evaluate the transcription in order to triangulate those themes I identified from the focus group.

**Summary**

Students in the online classroom were expected to have performed the same, if not better, than those in the traditional classroom, controlling for certain student characteristics. These results were to be validated in not just one semester, but over the ten semesters for which this study collected data. Moreover, the findings from the focus group were expected to reveal those perceptions that influence student satisfaction in the online version of Geol 105. Perhaps the results of the quantitative evidence and the revelations of the qualitative findings will help to strengthen our understanding of the effectiveness of the online environment in terms of student achievement, possibly impacting future online design strategies as well as improving faculty perceptions and attitudes to embrace such technologies in the classroom.
CHAPTER IV
QUANTITATIVE RESULTS

The results section is divided into two chapters. Chapter four analyzes the quantitative results followed by chapter five, a discussion of the qualitative findings. The quantitative research for this study used student data from both Geol 105 course environments – online and traditional face-to-face. For each student, the data files contained grades for test 1, test 2, final average in the course, incoming resident GPA the semester Geol 105 was taken, incoming overall GPA the semester Geol 105 was taken, number of previously taken online courses with a passing letter grade of D or better, declared degree program, number of completed student credit hours, course load for the semester, final letter grade, gender, ethnicity and date of birth. The traditional face-to-face Geol 105 data was from the Spring 2011, Spring 2012, Spring 2013, Spring 2014, and Spring 2015 semesters. The online Geol 105 data was from the Fall 2011, Fall 2012, Fall 2014, and Fall 2015 semesters.

Hypothesis One, Hypothesis Two, and Hypothesis Three

The first three hypotheses for this study were evaluated using ANCOVAs, which required scores for test 1, test 2, overall final average, final letter grade, and incoming GPA for both the online and the traditional, face-to-face Geol 105 courses. Because all students had entries for an incoming overall GPA while some had missing entries for incoming resident GPA, the incoming overall GPA was used as the covariate to evaluate the ANCOVAs.

**MatchIt Package in R.** Statistical studies that compare the mean between groups, such as ANCOVA, are considered more robust when the sample sizes of the two groups are equal
(Field, 2013). Therefore, the MatchIt package in the R programming language was used to do a propensity score matching (PSM) of the traditional face-to-face data based on mean (Randolph et al., 2014). The purpose of such an approach is to preserve the same percentage of data in each strata such that the mean of the reduced sample size is relatively close to that of the original sample size. For this research, letter grades were used as the strata, preserving the same overall percentages of As, Bs, Cs, Ds and Fs in the reduced Geol 105 face-to-face sample as in the complete data set.

Before running MatchIt, all five years of the Geol 105 face-to-face data was combined for test 1, test 2, the final score, incoming GPA, and the letter grade, which are the required fields for the first three hypotheses. Within this file, the percentages of As, Bs, Cs, Ds, and Fs was determined, as well as the corresponding mean and standard deviation for each grade. The original file with all the data was then parsed into five, one file for each strata’s data: A, B, C, D, and F. In addition to the student data, the five files for each strata – As, Bs, Cs, Ds, and Fs – also included an entry listing the means for that strata’s data set: test 1, test 2, the final score, and the overall incoming GPA. This additional record was required to match the treated group (the mean values) to the observed, control group (the remaining student data).

MatchIt was executed using a k-nearest neighbor approach to produce a reduced sample size of the data using a PSM approach. The sample size for the online class was N=171 so MatchIt was run to reduce the traditional face-to-face sample size to N=171. Table 2 lists the total number of As, Bs, Cs, Ds, and Fs in the original data set as well as the target number of As, Bs, Cs, Ds and Fs to be determined using MatchIt for the ANCOVA data set.
Table 2

*Percentage, Actual Sample Size, and Targeted Sample Size by Grade for the Face-to-Face Geol 105 Data*

<table>
<thead>
<tr>
<th>Letter Grade</th>
<th>Percentage</th>
<th>Actual Sample Size (N = 1266)</th>
<th>Reduced Target Sample Size - k (N = 171)</th>
</tr>
</thead>
<tbody>
<tr>
<td>As</td>
<td>53.3%</td>
<td>675</td>
<td>91</td>
</tr>
<tr>
<td>Bs</td>
<td>35.9%</td>
<td>455</td>
<td>61</td>
</tr>
<tr>
<td>Cs</td>
<td>8.7%</td>
<td>110</td>
<td>15</td>
</tr>
<tr>
<td>Ds</td>
<td>1.5%</td>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>Fs</td>
<td>0.6%</td>
<td>7</td>
<td>1</td>
</tr>
</tbody>
</table>

The approach of MatchIt is to match treated-group cases to control-group cases using a specified learning algorithm. When evaluating the various algorithms available for MatchIt, Randolph et al. (2014), found k-nearest neighbor to reduce the sample size most accurately when comparing to the original sample size. K-nearest neighbor is an instance-based learning algorithm that uses a clustering approach, finding cases that map closest to the mean in real space (Mitchell, 1997). The nearest “neighbors” are found by calculating the Euclidean distance from the center point based on the dimensions provided. In my research data, the dimensions are the three Geol 105 scores and the GPA.

For the Geol 105 data, the treated-group was made up of values, or nearest neighbor dimensions, containing the means from the two tests, the final score, and the GPA as shown in Table 3. The number of cases from the control-group, or the k in k-nearest neighbor, varied depending on the number required to preserve the corresponding percentage of As, Bs, Cs, Ds and Fs from the traditional Geol 105 class (Table 2). After running MatchIt and reducing the
data, one option in R is to list a summary report comparing the original control cases to the final observed cases, including both mean and standard deviation. This output allowed for confirmation that that the reduced sample size was a good representation of the original sample size, as shown in Table 4. However, because of assumption violations that were discovered in the data, which will be discussed later, another more robust ANCOVA was used that did not require equal sample sizes, ultimately making the use of MatchIt for PSM unnecessary.

Table 3

*Geol 105 Final Average, Exam Averages, and GPA Average by Grade from Complete Data Set*

<table>
<thead>
<tr>
<th>Letter Grade</th>
<th>Final Average</th>
<th>Exam 1</th>
<th>Exam 2</th>
<th>GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>91.96</td>
<td>90.00</td>
<td>89.75</td>
<td>3.30</td>
</tr>
<tr>
<td>B</td>
<td>83.58</td>
<td>82.25</td>
<td>82.10</td>
<td>2.68</td>
</tr>
<tr>
<td>C</td>
<td>74.63</td>
<td>72.87</td>
<td>74.15</td>
<td>2.27</td>
</tr>
<tr>
<td>D</td>
<td>65.63</td>
<td>69.23</td>
<td>58.88</td>
<td>2.42</td>
</tr>
<tr>
<td>F</td>
<td>40.09</td>
<td>40.19</td>
<td>29.94</td>
<td>1.45</td>
</tr>
</tbody>
</table>

Table 4

*Geol 105 Final Average, Exam Averages, and GPA Average by Grade from the MatchIt Data Set*

<table>
<thead>
<tr>
<th>Letter Grade</th>
<th>Final Average</th>
<th>Exam 1</th>
<th>Exam 2</th>
<th>GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>91.96</td>
<td>90.00</td>
<td>89.76</td>
<td>3.30</td>
</tr>
<tr>
<td>B</td>
<td>83.58</td>
<td>82.22</td>
<td>82.13</td>
<td>2.68</td>
</tr>
<tr>
<td>C</td>
<td>74.67</td>
<td>72.91</td>
<td>74.19</td>
<td>2.27</td>
</tr>
<tr>
<td>D</td>
<td>65.63</td>
<td>69.23</td>
<td>58.88</td>
<td>2.42</td>
</tr>
<tr>
<td>F</td>
<td>40.08</td>
<td>40.19</td>
<td>29.94</td>
<td>1.45</td>
</tr>
</tbody>
</table>
ANCOVA Assumptions. Before running an ANCOVA, it is important to consider the assumptions required for the data before analyzing and interpreting the results of the test. Disregard of these assumptions can lead to interpretations about the data that are not defensible (Field, 2013). For parametric data, there are four basic assumptions: (a) normally distributed data, (b) homogeneity of variance, (c) interval data, and (d) independence (Field, 2013).

The first of these assumption is normality. According to the Central Limit Theorem, sample sizes greater than 30 meet the assumption of normally distributed data (Field, 2013). Because the sample size was more than 30 for each course environment, N=171 for the online class and N=1266 for the face-to-face class, the data in this study easily met this first assumption. Additionally, sample sizes for individual semesters also met this assumption since each was also greater than 30 for each of the two environments. The second assumption, homogeneity of variance, requires that variance between groups is similar. In ANCOVA, this assumption is expanded to include the covariate and the consideration of homogeneity of regression, which will be discussed later. The third assumption, interval data, expects equal differences in data between each level. Because the data for this research considered performance in a Geol 105 class, which is based on a traditional grade distribution where A is 90% to 100%, B is 80% to 89%, C is 70% to 79%, D is 60% to 69%, and F is below 59%, the data set meets this assumption. The last general assumption is that of independence. For this study, students’ scores were independent because a student could not simultaneously enroll in both environments of Geol 105, and therefore the two groups were independent of each other.

In addition to the assumptions that must be met for analyzing parametric data, ANCOVA has two additional assumptions: (a) covariate and predictor variable independence, and (b) homogeneity of regression (Field, 2013). The first assumption is to verify that the independent
variable, or predictor variable, and the covariate are independent of each other. This is necessary because if they are dependent, then the statistic testing for the effect of the independent variable on the dependent variable would actually be a mix of the effects of both the independent variable and the covariate. In other words, it would be impossible to tell what percentage of the effect was due to just the independent variable since part of this effect would be from an overlap with the covariate, rendering an impossible interpretation of the independent variable’s sole effect on the dependent variable in the results.

In this study, meeting the first assumption required verifying independence between GPA (covariate) and course environment (independent variable) to show that there was no overlap in the effects of GPA and course environment on the Geol 105 score (dependent variable, or outcome). That is, for ANCOVA results to be meaningful, the variances in the outcome (Geol 105 scores) must be independently explained by each delivery mode of the class: GPA should not differ significantly by class environment. To verify independence between the covariate and independent variable, an ANOVA was run on SPSS comparing GPA and course environment.

<table>
<thead>
<tr>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>gpa</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>.026</td>
<td>1</td>
<td>.026</td>
<td>.059</td>
</tr>
<tr>
<td>Within Groups</td>
<td>644.556</td>
<td>1435</td>
<td>.449</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>644.582</td>
<td>1436</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 4. SPSS ANOVA output of GPA and Course Environment. ANOVA showing results of comparing GPA to course environment: online versus face-to-face. There is no significance in means between group (p = .809), therefore, GPA is independent of course environment.*
The result, shown in Figure 4, confirms there was no significant variance in GPA by course environment (p = .809), meeting the assumption of independence, which shows that by controlling for the effect of GPA, the significance of the course environment on the Geol 105 scores can be evaluated using ANCOVA.

The second assumption required for ANCOVA is homogeneity of regression, the slopes of the regression lines for each group (online and traditional face-to-face) should be similar. Because the ANCOVA analysis is trying to determine if the regression model is a better fit than the mean, the assumption is that the model is true for both groups, and therefore the regression lines for each are, more or less, parallel (Field, 2013). That is, the relationship between the covariate and the outcome are similar at each of the levels of GPA in both groups. In SPSS, this relationship can be determined by calculating the significance of the interaction between the covariate (GPA) and the independent variable (course environment). If the interaction is not significant, the two groups are behaving similarly at all levels and the assumption of the homogeneity of regression is met. Running three ANCOVAs in SPSS, for each Geol 105 score – final score, test 1, and test 2 – all showed a violation of the homogeneity of regression assumption, which would suggest an inflated Type I error, or incorrectly rejecting the null hypothesis when it actually is true (Hinkel et al., 2003).

Robust ANCOVA Using R. Because SPSS does not offer a robust ANCOVA calculation, as mentioned in chapter three, the programming language R was used instead to evaluate hypotheses one, two and three. R provides a robust ANCOVA package that uses trimmed means between groups, removing the requirement of distributional assumptions, including data sets that violate homogeneity of regression, or the second additional assumption for ANCOVA that has already been discussed (Field, Miles & Field, 2012). Furthermore, R’s
robust ANCOVA is not impacted by data samples that are unbalanced, as in the case of my data, N=171 for the online class and N=1266 for the face-to-face class (Field et al., 2012). Consequently, even though I initially used PSM to equalize the sample sizes for ANCOVA, this data reduction process proved to be unnecessary since R’s robust ANCOVA was used instead and therefore could use the complete sample sizes from both environments.

R’s robust ANCOVA does not require similar regression between data of equal sample sizes because this method identifies those five points from the covariate (GPA) where the two groups (course environment) have approximately the same relationship between the covariate and the outcome (Geol 105 scores), or approximately the same slopes at that instance (Field et al., 2012). For each of the five covariate points, the two groups’ trimmed means are calculated, using a 20% trimming factor, which is the typical value for trimmed means in a robust ANCOVA test (Field et al., 2012). To trim the mean, the percentage of the top and bottom outcome values (Geol 105 scores), or 20% in this case, are removed before calculating the mean. Once the trimmed mean is determined for each group at each of the five points, the corresponding differences are also computed, including the statistical significance of these differences. By using a robust ANCOVA test, the results show how differences in the group means (Geol 105 scores in each course environment) vary as a function of the covariate (GPA) (Field et al., 2012).

In addition to the robust ANCOVA, R also offers a robust bootstrap ANCOVA. The bootstrap procedure uses the sample itself as the sample population. A random sampling, known as a bootstrap sample, is used to calculate the mean. Several, often on the order of thousands, of bootstrap samples are taken, and the reported mean is the mean of all the bootstrap samples (Field, 2013). Because this sampling is random, each time a bootstrap sample is run, the
calculations will be slightly different. In R’s ANCOVA, the bootstrap-t is the same process as the robust ANCOVA, only a trimmed mean for the bootstrap samples is used rather than the original mean. For this data set, 2000 bootstrap samples were used, which is the typical number when implementing the bootstrap method (Field et al., 2012).

**Running R’s ancova and ancboot functions.** In order to determine the robust ANCOVA and robust bootstrap ANCOVA in R, the Wilcox Robust Statistics (WRS) package, which contains both the robust `ancova` and `ancboot` R functions, has to first be downloaded, since neither is included in the standard installation of R. The required file, Rallfun-v30.txt, was downloaded from Rand Wilcox’s website (http://dornsife.usc.edu/labs/rwilcox/software/), and then sourced into R (Windows platform): `source("C:/Rallfun-v30.txt")`. After downloading and installing the packages, both the `ancova` and `ancboot` functions were available. Both `ancova` and `ancboot` were executed for each of the three outcomes – final score, test 1, and test 2 – using GPA as a covariate. In all three cases, group 1 (n1) was the online environment and group 2 (n2) was the traditional face-to-face environment. Figure 5 shows the R summary output for the final scores using both of these robust ANCOVA tests – `ancova` and `ancboot`.

<table>
<thead>
<tr>
<th>X</th>
<th>Final Score</th>
<th>Test 1</th>
<th>Test 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA</td>
<td>DIF</td>
<td>GPA</td>
<td>DIF</td>
</tr>
<tr>
<td>2.0</td>
<td>0.0</td>
<td>2.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2.2</td>
<td>0.0</td>
<td>2.2</td>
<td>0.0</td>
</tr>
<tr>
<td>2.4</td>
<td>0.0</td>
<td>2.4</td>
<td>0.0</td>
</tr>
<tr>
<td>2.6</td>
<td>0.0</td>
<td>2.6</td>
<td>0.0</td>
</tr>
<tr>
<td>2.8</td>
<td>0.0</td>
<td>2.8</td>
<td>0.0</td>
</tr>
<tr>
<td>3.0</td>
<td>0.0</td>
<td>3.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

The X column lists the GPA values where GPA and final score were most closely related in both course environments. The n1 and n2 columns list the number of scores used from each group to calculate the trimmed mean. Based on the number of data points in each group, a 20% trimmed mean was calculated and the difference between these means is reported in the DIF column. For `ancova`, the standard error, se, is also given. Significance of this difference is reported in the p.value column. In addition to these results, the test statistic – mean difference between the two groups divided by the standard error – is given in the TEST column, along with the confidence interval lower bound and upper bound in the ci.low and ci.high columns,
respectively. As expected, the bootstrapping lower bound and upper bound from running
\textit{ancboot} differ from the \textit{ancova} evaluated without bootstrapping because it is based off of
bootstrap samples rather than the original sample.

\begin{verbatim}
[1] "NOTE: Confidence intervals are adjusted to control the probability"
[1] "of at least one Type I error."
[1] "But p-values are not"
$output$
   X  n1 n2        DIF     TEST
[1,] 1.49  18 186 -11.817607 3.437154
[2,] 2.51 105 722 -10.079444 8.797544
[3,] 3.01 132 900 -8.518641 8.972324
[4,] 3.43 108 836 -6.110729 6.447398
[5,] 4.00  57 489 -3.277441 3.067800
  se    ci.low      ci.hi
[1,] 3.4381951 -22.226424 -1.4087904
[2,] 1.1457111 -13.097369 -7.0615186
[3,] 0.9494353 -10.999137 -6.0381445
[4,] 0.9477822 -8.606480 -3.6149772
[5,] 1.0683361 -6.170404 -0.3844784
  p.value crit.val
[1,] 5.190634e-03 3.027407
[2,] 6.441514e-13 2.634107
[3,] 4.218847e-14 2.612602
[4,] 1.231894e-08 2.633254
[5,] 4.157749e-03 2.707915
[1] "Note: confidence intervals are adjusted to control FWE"
[1] "But p-values are not adjusted to control FWE"
\end{verbatim}
Figure 5. R output for robust *ancova* and *ancboot* functions. The output after running R’s *ancova* and *ancboot* functions where final score is the dependent variable, course environment the independent variable and GPA the covariate.

Besides the data results, R also generated a plot of the data, shown in Figure 6, which has the regression lines for each environment, online and face-to-face. In each of the subsequent R plots, the online environment is the solid line with asterisk (*) markers, and the face-to-face environment is the dashed line with plus (+) markers. The R^2 values for each of these regression lines has been included, which were determined plotting this same data in Microsoft Excel.

Figure 6. R plot of the final scores to GPA in the online and face-to-face data. R’s plot of the Geol 105 final scores (Y-axis) to GPA (X-axis) in both the online and face-to-face environments. The solid line with asterisks (*) – R^2 = .2971 – represents the online data. The dashed line with plusses (+) – R^2 = .2832 – represents the face-to-face data.
**Interpreting R’s Robust ANCOVA for Geol 105 Final Scores.** The column X in Figure 5 is the baseline GPA values where the relationship between GPAs and final scores were most comparable in each of the two environments: 1.49, 2.51, 3.01, 3.43, and 4.0. The next two columns, n1 and n2, list the number of data points (student final scores) used from each group to calculate the respective trimmed means. For example, of those students with a GPA of 1.49, 18 from the online Geol 105 students’ final scores (n1) was used to calculate the online trimmed mean, and 186 from the face-to-face Geol 105 students’ final scores (n2) was used to calculate the face-to-face trimmed mean. In other words, these students, 18 in the online Geol 105 (n1) and 186 in the face-to-face Geol 105 (n2), not only had close to a 1.49 GPA, but also had a similar relationship between their GPAs’ and their final scores’ in Geol 105.

In the second case, 105 online Geol 105 students’ final scores and 722 face-to-face Geol 105 students’ final scores were used to calculate the respective trimmed means because these students all had close to a 2.51 GPA. Similarly, the third case used 132 and 900 final scores from students in the online and face-to-face, respectively, because these students shared a GPA at, or near, 3.01. Likewise, final scores from 108 online students and 836 face-to-face students with close to 3.43 GPAs were used to compute the corresponding trimmed means. Finally, of those students with 4.0 GPAs, 57 from the online sections and 489 from the face-to-face sections were used for the trimmed mean calculations.

Based on these samples, a 20% trimmed mean was calculated for each group and the difference between these means is reported in the DIF column: -11.82, -10.08, -8.52, -6.11, and -3.28 for GPAs 1.49, 2.51, 3.01, 3.43, and 4.00, respectively. For ancova, the standard error, se, is also given. Significance of this difference is reported in the p.value column: .0035, .00, .00, .00, .006 for GPAs 1.49, 2.51, 3.01, 3.43, and 4.00, respectively. In addition to these results, the
test statistic – mean difference between the two groups divided by the standard error – is given in the TEST column, along with the confidence interval lower bound and upper bound in the ci.low and ci.high columns, respectively. As expected, the bootstrapping lower bound and upper bound from running \textit{ancboot} differ from the \textit{ancova} evaluated without bootstrapping because it is based off of random bootstrap samples rather than the original sample.

In all five instances of GPA, the difference in the means between group environments is significant, though the significance level is greater for GPAs between 2.51 and 3.43. This suggests there is a significant difference in final score means by course environment which aligns with some of the previous studies evaluating student performance in online and face-to-face classes (Atchley et al., 2013, Edmonds, 2006, Jaggars et al., 2013b). Because the difference is negative, (mean of online Geol 105 final score) minus (mean of face-to-face Geol 105 final score), students with similar GPAs enrolled in Geol 105 tend to have higher final scores in the traditional, face-to-face class than those in the online class. Moreover, this difference is more significant when a student’s GPA is between 2.51 and 3.43 (Field et al., 2012).

Because courses can differ not only by instructor, but also by semester, even with the same instructor teaching each of these semesters, as in the case of this research (i.e., the same online instructor taught all four semesters of online Geol 105 but was different than the face-to-face instructor, and the same face-to-face instructor taught all five semesters of the traditional Geol 105), the ANCOVA and bootstrap ANCOVA were also calculated for each year Geol 105 was offered: 2011, 2012, 2014, and 2015. The Spring 2011 face-to-face data was compared to the Fall 2011 online data, as well as Spring 2012 face-to-face to Fall 2012 online, Spring 2014 face-to-face to Fall 2014 online, and Spring 2015 face-to-face to Fall 2015 online. As mentioned earlier, the instructor for the online Fall 2013 course was a different instructor than the other four
semesters in which the online Geol 105 was taught, and was not used in the data for this research study. Because each of these courses in both environments had more than 30 students, the assumption of normality for ANCOVA was still satisfied. Table 5 lists the sample sizes used for each course environment for each year. Likewise, all of the remaining assumptions for ANCOVA were also met for the same reasons mentioned earlier, other than homogeneity of regression – running the interaction test in SPSS showed all four years of data violated the homogeneity of regression assumption. So, R was also used to evaluate the robust ANCOVA and robust bootstrap ANCOVA for each year.

Table 5

Course Environment Sample Sizes for Each Year of Data

<table>
<thead>
<tr>
<th>Year</th>
<th>Online Sample Size</th>
<th>Face-to-Face Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>40</td>
<td>163</td>
</tr>
<tr>
<td>2012</td>
<td>48</td>
<td>189</td>
</tr>
<tr>
<td>2014</td>
<td>41</td>
<td>352</td>
</tr>
<tr>
<td>2015</td>
<td>42</td>
<td>193</td>
</tr>
</tbody>
</table>

The complete R outputs for both `ancova` and `ancboot` are listed in Appendix C. A summary of these outputs, just the five GPAs used in each of the ANCOVAs, is listed in Table 6. The mean difference in student final scores for all cases was negative, meaning those students with similar GPAs in the face-to-face course outperformed those in the online course. However, not all of the differences were significant, p < 0.05, as shown by the starred (*) GPAs in Table 6.
Table 6

Summary of GPAs Used in R’s Robust ANCOVA for Geol 105 Final Scores

<table>
<thead>
<tr>
<th>GPAs</th>
<th>Complete Sample</th>
<th>2011</th>
<th>2012</th>
<th>2014</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.49</td>
<td></td>
<td>1.96</td>
<td>1.92</td>
<td>2.15*</td>
<td>2.27</td>
</tr>
<tr>
<td>2.51</td>
<td></td>
<td>2.36</td>
<td>2.48</td>
<td>2.65</td>
<td>2.80</td>
</tr>
<tr>
<td>3.01</td>
<td></td>
<td>2.65</td>
<td>2.90</td>
<td>3.10</td>
<td>3.33*</td>
</tr>
<tr>
<td>3.43</td>
<td></td>
<td>3.14*</td>
<td>3.35</td>
<td>3.36</td>
<td>3.57*</td>
</tr>
<tr>
<td>4.00</td>
<td></td>
<td>3.78*</td>
<td>3.92</td>
<td>3.91*</td>
<td>4.00*</td>
</tr>
</tbody>
</table>

Note. The negative differences in the five GPAs used to run the `ancova` R function are all significant, p < 0.05, except those that are starred (*).

Therefore, in considering both the complete data set and year-to-year data set, students with middle range GPAs tend to do better in the traditional face-to-face Geol 105 class than in the online Geol 105 class. In the case of this data set, middle range GPA would be a lower bound of 2.27, taken from 2015, to an upper bound of 3.10, taken from 2014. Other than 2014, the data suggests that even lower performing students (GPAs less than 2.27) also do significantly worse in the online Geol 105 than in the face-to-face Geol 105, but this evaluation is not as conclusive as the middle-range GPAs because of the 2014 results. Additionally, the ANCOVA results suggest that if a student has a history of being a high-performing student, indicative by a high GPA, the course environment does not have a significant impact on their overall performance in the course.

R’s Robust ANCOVA for Geol 105 for Test 1 and Test 2. The same approach for the final scores was also used to evaluate the first and second tests in Geol 105. As in the case of the final score, the data did not meet the required homogeneity of regression assumption, neither for
the complete data set nor for the year-to-year data sets. Consequently, R was used to evaluate both the robust ANCOVA and robust bootstrap ANCOVA for test 1 and test 2. Complete results of the R outputs are listed in Appendix C. A summary of these results are listed in Table 7 for test 1 and Table 8 for test 2. Not surprising, the results showed similar conclusions as to what was determined for the final score, students in Geol 105 with middle range GPAs tend to do better in a face-to-face environment than in an online environment. Also, high-performing students seem to do well regardless of the environment.

Table 7

Summary of GPAs Used in R’s Robust ANCOVA for Geol 105 Test 1

<table>
<thead>
<tr>
<th>GPAs</th>
<th>Complete Sample</th>
<th>2011</th>
<th>2012</th>
<th>2014</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.49</td>
<td>1.96</td>
<td>1.92</td>
<td>2.15</td>
<td>2.27</td>
<td></td>
</tr>
<tr>
<td>2.51</td>
<td>2.36</td>
<td>2.48</td>
<td>2.65</td>
<td>2.80</td>
<td></td>
</tr>
<tr>
<td>3.01</td>
<td>2.65</td>
<td>2.90</td>
<td>3.10</td>
<td>3.33</td>
<td></td>
</tr>
<tr>
<td>3.43</td>
<td>3.14</td>
<td>3.35</td>
<td>3.36</td>
<td>3.57*</td>
<td></td>
</tr>
<tr>
<td>4.00</td>
<td>3.78</td>
<td>3.92</td>
<td>3.91*</td>
<td>4.00*</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* The negative differences in the five GPAs used to run the *ancova* R function are all significant, $p < 0.05$, except those that are starred (*).
Table 8

*Summary of GPAs Used in R’s Robust ANCOVA for Geol 105 Test 2*

<table>
<thead>
<tr>
<th>GPAs</th>
<th>Complete Sample</th>
<th>2011</th>
<th>2012</th>
<th>2014</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.49</td>
<td></td>
<td>1.96*</td>
<td>1.92*</td>
<td>2.15</td>
<td>2.27</td>
</tr>
<tr>
<td>2.51</td>
<td></td>
<td>2.36*</td>
<td>2.48*</td>
<td>2.65</td>
<td>2.80</td>
</tr>
<tr>
<td>3.01</td>
<td></td>
<td>2.65</td>
<td>2.90</td>
<td>3.10</td>
<td>3.33*</td>
</tr>
<tr>
<td>3.43</td>
<td></td>
<td>3.14*</td>
<td>3.35</td>
<td>3.36</td>
<td>3.57*</td>
</tr>
<tr>
<td>4.00</td>
<td></td>
<td>3.78*</td>
<td>3.92</td>
<td>3.91*</td>
<td>4.00*</td>
</tr>
</tbody>
</table>

*Note.* The negative differences in the five GPAs used to run the *ancova* R function are all significant, *p < 0.05,* except those that are starred (*).

**Hypothesis Four**

The fourth hypothesis in this study considered possible variables for predicting student performance in the online Geol 105 class: incoming GPA, number of prior online courses completed with a D or better, type of declared degree program – STEM or non-STEM, Geology or non-Geology – number of completed student credit hours, and current course load. In SPSS, these predictor variables were used to determine an OLS regression equation for calculating a student’s expected performance in the online Geol 105 course.

The data for the predictor variables provided by Institutional Research, Effectiveness and Planning (IREP) included two GPAs, incoming resident GPA and incoming overall GPA. Because some data was missing from incoming resident GPA, but not from incoming overall GPA, as was the case for hypotheses one, two, and three, the incoming overall GPA was used for the OLS analysis.
In order to determine STEM v. non-STEM degree programs and Geology v. non-Geology degree programs, the original data had to be modified for SPSS, converting the data into two columns of ordinal data instead: 1 or 0 for STEM or non-STEM major, respectively, in the first column, and 1 or 0 for Geology or non-Geology major, respectively, in the second column. The Excel file containing all of the original online Geol 105 data was sorted by degree program, and those majors that were STEM had a 1 entered in the STEM column: Biology, Chemistry, Math Education, Science Education, General Engineering, Geology, Civil Engineering, Computer Science, and Geological Engineering. In addition, those students with a declared degree program of Geology or Geological Engineering had a 1 entered in the Geology column. If students did not fall into either of these respective categories, a 0 was entered into the columns.

Although hypothesis four only conjectured the relationship for those students in the online course, this same process of modifying the data was done for the face-to-face Geol 105 course. By determining a regression equation for each class environment, findings from hypothesis three using the robust ANCOVA for the final score could be further investigated and will be discussed later.

**OLS Regression Assumptions.** Like ANCOVA, OLS also has assumptions that must be met regarding a linear model. Failure to meet these assumptions confounds the interpretation of the results (Field, 2013). Regression has four assumptions: (a) linearity, (b) homogeneity of variance, (c) independent errors, and (d) normality (Field, 2013).

The first assumption is linearity – the combination of predictor variables should linearly relate to the outcome (Field, 2013). For this study, linearity means the best model for predicting all final scores in the online Geol 105 course is to take the sum of each of the predictor variables for an observed set of predictor variables multiplied by their respective coefficients. The second
assumption is homogeneity of variance, or equal variance among residuals at all levels of the final scores. A residual term is the difference between the model predicting the outcome and the actual outcome observed. In the case of the first two assumptions, both can be verified using a scatterplot of the residual values relative to the predicted outcome of the model (final score), which was generated in SPSS and shown in Figure 7.

Figure 7. Standardized residuals of Geol 105 final scores to standardized predicted values. The scatterplot of the standardized residuals of Geol 105 final scores to the standardized predicted values was used to verify the first two assumptions required for OLS regression.

For the first two assumptions to be met, the plot should not funnel or curve, which it does not. Therefore, the data meets the first two assumptions. In looking at the standardized residual axis, there are seven points that are more than 2 standard deviations from the mean (-2 to -5). Because these points make up less than 5% of the data (i.e., 7 out of 171, or 4.1%), for a confidence interval of 95%, this few number of outlier points still meets the requirement that no more than 5% of the data exceed two standard deviations (Field, 2013).

The third assumption relates to independence of error terms in that each of the data points’ errors should be independent of the other data point errors. The Durbin-Watson test can be used to verify this assumption, which evaluates correlations between errors (Field, 2013).
the test statistic is close to 2 – it ranges from 0 to 4 – then this assumption is met. Figure 8 shows the SPSS summary, including the result of this test. Because the Durbin-Watson is close to 2, 2.068 to be exact, the independence assumption is met.

The fourth assumption is normalized data. Normality was discussed in the section on the first three hypotheses – data with more than 30 points is normalized according to the Central Limit Theorem (Field, 2013). Because this research used data from 171 students, the data set meets this fourth assumption. Additionally, data from each of the semesters for this study had more than 30 students and therefore also met this assumption.

**Model Summary**

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
<th>R Square Change</th>
<th>F Change</th>
<th>df1</th>
<th>df2</th>
<th>Sig. F Change</th>
<th>Durbin-Watson</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.545a</td>
<td>.297</td>
<td>.293</td>
<td>13.57513</td>
<td>.297</td>
<td>71.427</td>
<td>1</td>
<td>169</td>
<td>.000</td>
<td>2.068</td>
</tr>
<tr>
<td>2</td>
<td>.556b</td>
<td>.309</td>
<td>.301</td>
<td>13.50136</td>
<td>.012</td>
<td>2.852</td>
<td>1</td>
<td>168</td>
<td>.093</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.573c</td>
<td>.328</td>
<td>.316</td>
<td>13.34976</td>
<td>.019</td>
<td>4.837</td>
<td>1</td>
<td>167</td>
<td>.029</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.577d</td>
<td>.333</td>
<td>.317</td>
<td>13.34435</td>
<td>.005</td>
<td>1.136</td>
<td>1</td>
<td>166</td>
<td>.288</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.579e</td>
<td>.335</td>
<td>.315</td>
<td>13.36097</td>
<td>.002</td>
<td>.587</td>
<td>1</td>
<td>165</td>
<td>.445</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.579f</td>
<td>.335</td>
<td>.311</td>
<td>13.40162</td>
<td>.000</td>
<td>.000</td>
<td>1</td>
<td>164</td>
<td>.984</td>
<td></td>
</tr>
</tbody>
</table>

a. Predictors: (Constant), gpa  
b. Predictors: (Constant), gpa, courseLoad  
c. Predictors: (Constant), gpa, courseLoad, SCH  
d. Predictors: (Constant), gpa, courseLoad, SCH, STEM  
e. Predictors: (Constant), gpa, courseLoad, SCH, STEM, Geol  
f. Predictors: (Constant), gpa, courseLoad, SCH, STEM, Geol, onlineCourses  
g. Dependent Variable: final

*Figure 8.* SPSS model summary output showing $R^2$ and the Durbin-Watson test statistic.
Multiple regression has three added assumptions: (a) all predictor variables must be continuous or categorical, (b) values of predictors should have variation, and (c) multicollinearity.

<table>
<thead>
<tr>
<th>Model</th>
<th>Excluded Variables</th>
<th>Beta In</th>
<th>t</th>
<th>Sig.</th>
<th>Partial Correlation</th>
<th>Collinearity Statistics</th>
<th>Minimum Tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>courseLoad</td>
<td>.110b</td>
<td>1.689</td>
<td>.093</td>
<td>.129</td>
<td>.970</td>
<td>1.031</td>
</tr>
<tr>
<td></td>
<td>SCH</td>
<td>.105b</td>
<td>1.583</td>
<td>.115</td>
<td>.121</td>
<td>.928</td>
<td>1.078</td>
</tr>
<tr>
<td></td>
<td>STEM</td>
<td>.094b</td>
<td>1.456</td>
<td>.147</td>
<td>.112</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Geol</td>
<td>.109b</td>
<td>1.693</td>
<td>.092</td>
<td>.130</td>
<td>.985</td>
<td>1.016</td>
</tr>
<tr>
<td></td>
<td>onlineCourses</td>
<td>.010b</td>
<td>.147</td>
<td>.883</td>
<td>.011</td>
<td>.993</td>
<td>1.007</td>
</tr>
<tr>
<td>2</td>
<td>SCH</td>
<td>.151c</td>
<td>2.199</td>
<td>.029</td>
<td>.168</td>
<td>.848</td>
<td>1.179</td>
</tr>
<tr>
<td></td>
<td>STEM</td>
<td>.087c</td>
<td>1.360</td>
<td>.176</td>
<td>.105</td>
<td>.996</td>
<td>1.004</td>
</tr>
<tr>
<td></td>
<td>Geol</td>
<td>.100c</td>
<td>1.554</td>
<td>.122</td>
<td>.119</td>
<td>.977</td>
<td>1.024</td>
</tr>
<tr>
<td></td>
<td>onlineCourses</td>
<td>.031c</td>
<td>.477</td>
<td>.634</td>
<td>.037</td>
<td>.958</td>
<td>1.044</td>
</tr>
<tr>
<td>3</td>
<td>STEM</td>
<td>.068d</td>
<td>1.066</td>
<td>.288</td>
<td>.082</td>
<td>.975</td>
<td>1.025</td>
</tr>
<tr>
<td></td>
<td>Geol</td>
<td>.081d</td>
<td>1.250</td>
<td>.213</td>
<td>.097</td>
<td>.954</td>
<td>1.048</td>
</tr>
<tr>
<td></td>
<td>onlineCourses</td>
<td>-.009d</td>
<td>-.138</td>
<td>.890</td>
<td>-.011</td>
<td>.883</td>
<td>1.133</td>
</tr>
<tr>
<td>4</td>
<td>Geol</td>
<td>.062e</td>
<td>.766</td>
<td>.445</td>
<td>.060</td>
<td>.623</td>
<td>1.606</td>
</tr>
<tr>
<td></td>
<td>onlineCourses</td>
<td>.000e</td>
<td>.006</td>
<td>.995</td>
<td>.000</td>
<td>.866</td>
<td>1.154</td>
</tr>
<tr>
<td>5</td>
<td>onlineCourses</td>
<td>.001f</td>
<td>.020</td>
<td>.984</td>
<td>.002</td>
<td>.866</td>
<td>1.154</td>
</tr>
</tbody>
</table>

a. Dependent Variable: final
b. Predictors in the Model: (Constant), gpa
c. Predictors in the Model: (Constant), gpa, courseLoad
d. Predictors in the Model: (Constant), gpa, courseLoad, SCH
e. Predictors in the Model: (Constant), gpa, courseLoad, SCH, STEM
f. Predictors in the Model: (Constant), gpa, courseLoad, SCH, STEM, Geol

Figure 9. SPSS excluded variables output for linear regression, which lists the variance inflation factor (VIF), to determine if multicollinearity exists.
should not exist among the predictors (Field, 2013). The first two assumptions are self-evident in the data. The third, however, required analysis that can be done in SPSS. Multicollinearity occurs when two predictors are highly correlated. The result is that the regression coefficients include overlap of the combination of these predictors on the outcome, making it impossible to ascertain the effect each one alone has on the outcome. In SPSS, collinearity statistics provide the variance inflation factor (VIF), which can be used to determine if multicollinearity exists. As long as the VIF values are less than 10, and the average VIFs are not significantly greater than 1, then none of the variables are correlated (Field, 2013). Figure 9 shows the SPSS output for the possible multiple regression models. None of the VIF values are close to 10, and none of the average VIF values are substantially greater than 1, so the assumption is met.

**SPSS Multiple Regression.** In order to ascertain the influence of each predictor variable on the outcome, a hierarchical regression approach was used in SPSS, introducing a new

Table 9

*Predictor Variables Used in Each Stage of Regression Analysis*

<table>
<thead>
<tr>
<th>Model</th>
<th>Predictor Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GPA</td>
</tr>
<tr>
<td>2</td>
<td>GPA, course load</td>
</tr>
<tr>
<td>3</td>
<td>GPA, course load, student credit hours</td>
</tr>
<tr>
<td>4</td>
<td>GPA, course load, student credit hours, STEM/non-STEM major</td>
</tr>
<tr>
<td>5</td>
<td>GPA, course load, student credit hours, STEM/non-STEM major, Geol/non-Geol major</td>
</tr>
<tr>
<td>6</td>
<td>GPA, course load, student credit hours, STEM/non-STEM major, Geol/non-Geol major, number of online courses scoring a D or better</td>
</tr>
</tbody>
</table>
predictor variable for each subsequent model. A list of these models and the corresponding predictor variables are listed in Table 9. The first model only used GPA as a predictor with subsequent stages adding one additional predictor variable each time, up to the last model, which included all six predictor variables for this study. The order in which to add predictor variables was determined by previous research (Driscoll et. al, 2012, Hachey et al., 2014).

The SPSS output that was shown in Figure 8 lists the significance of each of the six models. Only models one and three were significant: GPA (p < .001), and the combination of GPA, course load and student credit hours, respectively (p = .029). In order to verify that GPA in combination with just one other predictor variable was not another possible significant model, SPSS tests were also run with just 2 models: GPA and the combination of GPA and just one of the other predictor variables. In all cases, none of the second models was significant.

From the SPSS model summary output (Figure 8), the $R^2$ values list the percentage that can be explained by each of the predictor variables. GPA explains 29.7% of the final score outcome, course load explains 1.2% (i.e., .309 minus .297) of the final score outcome, and student credit hours explains 1.9% (i.e., .328 minus .309) of the final score outcome.

The last step in multiple regression is to determine the OLS regression equation for each of the significant models that predict performance in an online Geol 105 class. Figure 10 lists the SPSS output that includes the constants and coefficients for models one and three. These findings are also summarized in Table 10. The regression equation for model 1 is as follows:

$$Geol\ 105\ final\ score_i = 33.244 + (14.895 \times GPA_i)$$

The first term is the constant, followed by the second term, the product of the GPA predictor variable and its coefficient, with a significance of p < .001 (Figure 8). For model 3, the regression equation is as follows:
Geol 105 final score\(_i\) = 17.223 + (15.279 \times \text{GPA}_i) + (0.727 \times \text{courses}_i) + (0.061 \times \text{SCH}_i)

Because this latter equation had three predictor variables instead of just one, there are two additional terms in the OLS regression equation: the product of the number of courses and its coefficient as well as the product of the number of student credit hours and its coefficient. The significance for model 3 is \(p = .029\) (Figure 8).

In addition to using data that is combined for all semesters, the same analysis was conducted for each year: 2011, 2012, 2014, and 2015. Table 11 lists the models by year that were significant (\(p < .05\)).

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>95.0% Confidence Interval for B</th>
<th>Correlations</th>
<th>Collinearity Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td>t</td>
<td>Sig.</td>
</tr>
<tr>
<td>1</td>
<td>(Constant)</td>
<td>33.244</td>
<td>5.305</td>
<td>.6267</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Gpa</td>
<td>14.895</td>
<td>1.762</td>
<td>.545</td>
<td>.000</td>
</tr>
<tr>
<td>3</td>
<td>(Constant)</td>
<td>17.223</td>
<td>7.780</td>
<td>.214</td>
<td>.028</td>
</tr>
<tr>
<td></td>
<td>Gpa</td>
<td>15.279</td>
<td>1.807</td>
<td>.559</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>courseLoad</td>
<td>.727</td>
<td>.319</td>
<td>.153</td>
<td>.227</td>
</tr>
<tr>
<td></td>
<td>SCH</td>
<td>.061</td>
<td>.028</td>
<td>.151</td>
<td>2.199</td>
</tr>
</tbody>
</table>

a. Dependent Variable: final

Figure 10. SPSS coefficients output. Output is for the model 1’s and model 3’s coefficients, the only two significant models of the six models.
### Table 10

**Summary of the SPSS Coefficients Output**

<table>
<thead>
<tr>
<th>Model</th>
<th>b, 95% CI</th>
<th>SE B</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>33.244, 5.305</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[22.772, 43.717]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPA</td>
<td>14.895, 1.762</td>
<td></td>
<td>.545</td>
</tr>
<tr>
<td></td>
<td>[11.416, 18.375]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Model 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>17.223, 7.780</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[1.864, 32.582]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPA</td>
<td>15.279, 1.807</td>
<td></td>
<td>.559</td>
</tr>
<tr>
<td></td>
<td>[11.711, 18.847]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Course Load</td>
<td>.727, 0.153</td>
<td></td>
<td>0.153</td>
</tr>
<tr>
<td></td>
<td>[.097, 1.357]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCH</td>
<td>0.061, 0.151</td>
<td></td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>[.006, .115]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: The information listed only includes the necessary information for the OLS regression equation.*
Table 11

*Summary of the SPSS Coefficients Output by Year*

<table>
<thead>
<tr>
<th>Year</th>
<th>Model 1</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2011, (N = 40)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Constant</td>
<td>Constant</td>
<td>Constant</td>
</tr>
<tr>
<td></td>
<td>GPA</td>
<td>GPA</td>
<td>GPA</td>
</tr>
<tr>
<td></td>
<td>β</td>
<td>β</td>
<td>β</td>
</tr>
<tr>
<td></td>
<td>.533</td>
<td>.485</td>
<td>.592</td>
</tr>
<tr>
<td>Fall 2012, (N = 48)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Constant</td>
<td>Constant</td>
<td>Constant</td>
</tr>
<tr>
<td></td>
<td>GPA</td>
<td>GPA</td>
<td>GPA</td>
</tr>
<tr>
<td></td>
<td>β</td>
<td>β</td>
<td>β</td>
</tr>
<tr>
<td></td>
<td>.496</td>
<td>.485</td>
<td>.612</td>
</tr>
</tbody>
</table>
Table 11 continued

<table>
<thead>
<tr>
<th>Year</th>
<th>b, 95% CI</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course Load</td>
<td>0.939, [-0.483, 2.360]</td>
<td>.181</td>
</tr>
<tr>
<td>SCH</td>
<td>0.032, [-0.104, 0.169]</td>
<td>.067</td>
</tr>
<tr>
<td>STEM</td>
<td>-8.099, [-24.266, 8.068]</td>
<td>-.129</td>
</tr>
<tr>
<td>Geol</td>
<td>-52.505, [-89.781, -15.229]</td>
<td>-.392</td>
</tr>
</tbody>
</table>

Fall 2014, Model 1 (N = 41)

| Constant | 39.623, [19.668, 59.578] |

Fall 2015, Model 1 (N = 42)

| Constant | 41.366, [22.734, 59.998] |
| GPA      | 13.291, [7.452, 19.130]  |

Note: The information listed only includes the necessary information for the OLS regression equation.

Summary

For the ANCOVA analyses, classroom environment does have influence on student outcome in most cases, after controlling for GPA. The one instance where environment does not have as much of an impact is for the strongest performing students, those with GPAs above 3.9. Moreover, this conclusion was verified when determining the possible predictor variables for performance in Geol 105. As shown in hypothesis four, the only predictor variables with
significance on final score in the online Geol 105 course were GPA, the number of courses a student was taking in the same semester as Geol 105, and the year they were in school – Freshman, Sophomore, Junior, Senior – as determined by the number of student credit hours completed. However, of these three variables, GPA had a much stronger effect, more than seven times that of course load and SCH combined (i.e., GPA had 29.7% effect on the outcome compared to course load and SCH which had a combined effect of 4.1%). Therefore, considering just GPA as a predictor for performance in Geol 105, the conclusion of hypothesis four validates that of the ANCOVA results, and classroom environment does significantly influence student performance.
CHAPTER V

QUALITATIVE FINDINGS

The qualitative research for this study used findings from a focus group conducted with four students enrolled in the online Geol 105 Spring 2016 semester. The purpose of the focus group was to explore those characteristics that contributed to satisfaction in courses conducted in the online environment, specifically Geol 105.

Subject Selection

Initially, the plan was to have as many as twelve student volunteers for the focus group to discuss those issues related to their experiences in the online Geol 105 course. Moreover, the expectation of the study was to have enough volunteers such that a diverse group could be selected by current grade in the class. Therefore, after midterm grades were posted for the university, the online Geol 105 instructor posted an announcement and sent out an email to the students asking for volunteers for this research study. The email not only described the objectives of the focus group, but also highlighted that participation was not linked to their grade in the course in any way. The email also included information regarding the added incentive of a gift card for all who volunteered. Students were asked to directly email the moderator if they were interested. All email interchanges are listed in Appendix A.

After the instructor sent out the first email, only three students replied that they were interested in participating in the study – two female students and one male student. Consequently, a week later, the instructor sent out a second email, again asking for volunteers.
Following this second attempt, only one more student volunteered, a male student, who, like the other three from the first email, replied the first day the email was sent out. At this point, a decision had to be made about whether to proceed with the focus group or to try and solicit more volunteers.

Although the preferred number of individuals for a focus group has been identified as seven to ten, it is not necessarily a critical mass as the primary purpose of the focus group is a meaningful discussion of the topic (Krueger & Casey, 2000). Vaughn et al. (1996) stated the exact number of participants is less important than creating a stimulating dialogue, giving credibility to the findings. Moreover, the procedure for selecting the focus group sample is through purposive sampling, the size of the group is less important than the characteristics of the participants, and researchers recommend a homogenous group over a heterogeneous one (Vaughn et al., 1996). In this research, volunteers for the discussion (a) needed to be enrolled in the Spring 2016 online Geol 105, and (b) had to be willing to share their experiences taking this course in an online environment such that their responses were believable and trustworthy (Vaughn et al., 1996). A common problem sometimes identified with purposive sampling is generalizability. However, as in the case of this research study, generalizing findings from the focus group to a larger population was not the objective but rather to explore how these particular students described their experiences and satisfaction in an online class. As discussed later in possible future work, conducting more focus groups related to online education with students enrolled in online classes could serve to strengthen the findings of this and previous studies (Armstrong, 2011; Carter & Emerson, 2012; Cole et al., 2014; Kuo et al., 2013), but determining how many focus groups “is based on the purpose of the study, the background information the individual researcher needs, the nature of the focus group, and the success of the
first focus group” (Vaughn et al., 1996, p. 48). As will be discussed in the analysis, the focus
group dialogue was lively and informative, meeting the objectives of a successful focus group.

Therefore, after the second email interchange with students in the Spring 2016 online
Geol 105 class, the decision was made to move forward with the focus group since four was a
reasonable number of participants with whom to have a discussion, and there was little
expectation that more than these four students would volunteer, even with additional emails to
the class. Although the group was not diverse by grade in the course, they did meet the
requirements of purposive sampling, including homogenous characteristics. Thus, the next step
was taken in the focus group protocol of identifying a time and place of where to conduct the
focus group.

Planning and Conducting the Focus Group

An email was sent out to the four students, Kelsie, Haley, Josh and Ryan, letting them
know that the focus group would be held the week of April 15 but that they needed to reply what
afternoons they were available to meet (see Appendix A). Surprisingly, finding a common time
for everyone was easy and only required two email interchanges in order to establish a date and
time.

The physical setting is an important factor when scheduling a focus group (Vaughn et al.,
1996). When choosing the focus group location, the researcher should consider the size of the
room, one that is neither too large nor too small, creating an inviting atmosphere for dialogue.
The room should also allow for easy recording of the discussion and should be in a location that
is easily accessible by the participants (Vaughn et al., 1996). Although many focus groups are
held in locations that are viewed as a retreat, or a get-away from individuals’ normal routines, it
was decided by the researcher that these students were more likely to participate if the focus
group was conducted on the main campus since three of the four students worked on campus and the fourth lived on campus (Vaughn et al., 1996). Consequently, the location of the focus group was to be conducted in a medium-sized conference room on the university’s main campus. A final email was sent notifying the four that the focus group was scheduled to be conducted from 3pm to 4:30pm on Wednesday, April 20 in this conference room. As anticipated, all four respondents preferred the focus group to be conducted on the main campus. The students were asked to plan for an hour and a half focus group, even though the researcher’s intent was to have just an hour discussion. By asking for this extra time, the hope was that the four would be more inclined to fully participate the entire hour, even if it went over the planned sixty minutes. Additionally, the questions to discuss were emailed to the four volunteers ahead of time with the qualification that the focus group was not limited to these questions. The rationale for communicating this information upfront was for students to know the intent of the questions before their answers would be recorded in a discussion, as well as to allow them time to consider their answers before the focus group was actually conducted.

Even though it is often the case that at least one participant will not show, or show up late, everyone arrived at the conference room at the requested time (Vaughn et al., 1996). Before recording the interview, as the moderator, I began by giving the participants their gift cards. I then reiterated the information in the email explaining that the focus group would be recorded, that we would begin with introductions, and, most importantly, that they be honest in their input regarding their experiences in the online Geol 105 class. The group was made up of one female Freshman Accounting major, Kelsie Smith, one male Sophomore Computer Science major, Ryan Jones, and two Juniors, a male Computer Science major, Josh Williams, and a female Integrated Marketing Communications major, Haley Adams. All four students volunteered they were
currently earning either an A or a B in the course. Because I knew the two male students ahead of time, I was confident they would participate. I was also pleasantly surprised that two of the students, Haley and Josh, remembered each other from a class they had together when they were freshmen. So, before the focus group began, my expectation was that, even though the group was small in size, there would be good discussion among everyone, and the focus group would be a success.

The focus group lasted one hour, and, as hoped, all four students fully participated. Even though she was the youngest in the group, Kelsie was inclined to dominate the conversation, so I would purposely direct questions to Josh, Ryan and Haley to keep them active in the discussion. Additionally, about halfway into the focus group time, Josh and Ryan were getting distracted looking at their cell phones, and I quickly realized it was important to actively draw them into the discussion. Specifically, when Josh first looked at his cell phone, to regain his attention, I asked what he thought of the question – *overall, how satisfied are you with online Geol 105*. At that point, he asked what the question was, which clued me in that when he was looking at his cell phone he was completely disengaged from the group discussion. From that point on, I was mindful of specifically requesting input from him to keep him involved in the focus group.

Another interesting observation was related to Ryan. He too had great input, but when he saw Josh check his messages on his cell phone, he soon did the same. So, for Josh and Ryan, I tried to intentionally ask for their input to most of the remaining questions so they would not look at their cell phones. Neither Kelsie nor Haley looked at their cell phones. Overall, I felt that there was excellent discussion among the four, with little to no straying from the topic of their impressions of online classes.
After an hour, we had covered the topics I had hoped to discuss (Table 1) and conversation was beginning to re-visit issues we had already addressed. Specifically, the focus group began with topic of convenience and, after covering all of my questions and more, when conversation returned to the topic of convenience, I felt we had saturated the discussion of their experiences in the online class and therefore concluded the focus group.

Emerging Themes

The recorded online Geol 105 focus group was transcribed, and NVivo 11CAQDAS software was used to help organize the coding of the data. Initially themes identified in Cole et al.’s (2014) research were used as a starting point for analysis: convenience, structure, learning style, interaction, and learning platform. Overall, the data confirmed these findings, but as will be discussed, some themes required some modification. Once the data was coded, two other individuals were asked to triangulate the findings. The purpose of investigator triangulation is to test for the consistency of the findings (Patton, 2002). The first person asked to code the focus group is a specialist in online instructional design. The second person, who also transcribed the data, is a medical student, experienced in coding other focus groups for research conducted at the university. All three coders concurred in our findings of those themes that emerged in the focus group.

Interaction. While Cole et al. (2014) found student interaction with the instructor to be a prominent theme, this focus group had the theme of interaction encompassing all three tenets of Moore’s (1993) transactional distance theory: student-to-instructor, student-to-student, and student-to-content, which is similar to the findings of Kuo et al. (2013).

Student-to-Instructor Interaction. All four students commented that the instructor for Geol 105 was good about communicating expectations upfront, specifically emailing the class
that to be successful in an online course, students must be organized and manage their time throughout the semester. Other than this first email, however, the only regular communication was through announcements posted on Blackboard, the learning management system used to conduct the course. These announcements were also emailed to students, but an interesting observation was that the students could not recall the content of what was sent. As Haley pointed out, the subject line started with “Do Not Reply” (the do not reply is controlled by the Blackboard administrator, not the instructor). This seemed to have a negative connection with the students. As Josh stated, “honestly, up until this, I don’t think I knew my geology professor’s name” (J. Williams, personal communication, April 20, 2016). All four definitively expressed that receiving personal emails would have helped to create a better connection with the instructor rather than just receiving email announcements.

Only one of the four, Kelsie, had initiated a few email interchanges with the instructor, and she said she always received prompt email replies from the instructor. Kelsie specifically recalled an online quiz in which she had technical issues. She commented that the instructor quickly responded to her concern, rectifying the issue, which, as she stated, “was nice to know that I wasn’t so left out in the open if something happens or something happens with technology and you just fail the quiz” (K. Smith, personal communication, April 20, 2016).

**Student-to-Student Interaction.** In addition to interaction with the instructor, students in the focus group all expressed a desire to have some connection with other students in the class, even though the course was conducted remotely. Ryan mentioned he wished there was opportunity to email other students in the class, but he did not know how to do this. He could recall from other classes he had taken that “I remember being able to email other kids in my class
if I didn’t know their email. There was a way for me to find the email in like the roster or something inside Blackboard” (R. Jones, personal communication, April 20, 2016).

Additionally, when discussing ways of improving online Geol 105, Haley mentioned it would be nice to have the instructor schedule time in a reserved room each week like in one of her previous online classes, where students in the course could go and talk to either the instructor or other students in the course. She praised this approach because you “see people you know, meet new people, and actually makes you more excited about taking the course. I don’t know why but when I saw three people I knew, I was like, this is so cool” (K. Smith, personal communication, April 20, 2016).

**Student-to-Content.** Convenience is undoubtedly the primary reason these four students enrolled in Geol 105, but they all agreed that the structure of the face-to-face classroom would have helped them to learn the content better. Ryan encapsulated this finding when he stated, “I mean I like the convenience of [the online environment], like not having to be in the classroom, but I feel like the structure of the classroom kind of gives you more drive to actually do the information and kind of get more out of it” (R. Jones, personal communication, April 20, 2016). Haley’s suggestion for promoting education in an online environment was for instructors to offer in-person office hours for those students desiring to learn the material. In a previous online class she had taken, she mentioned going every week to her online professor’s office hours to discuss the content, “With the online class I took in my major, the reason I actually retained everything from that is because the teacher who was doing that online course was really available in the building over on campus, and so I would go and talk to him a lot about the lesson we did. So, I got the visual I needed, but I didn’t have to go to the class every week” (H. Adams, personal communication, April 20, 2016). It was evident by her description of this online course that she
significantly enjoyed it more than her Geol 105 course, and it was, in part, because she could learn by both online material and face-to-face meetings with the instructor.

While Geol 105 lessons included PowerPoint slides with audio, outside reading assignments, and YouTube videos, all but Kelsie admitted to devoting only minimal time to the content. Specifically, Josh acknowledged, “I don’t watch her videos. If it was a class like a major class, I would absolutely love online videos because I could rewind and take great notes” (J. Williams, personal communication, April 20, 2016). Ryan suggested that incorporating some type of midweek checkpoints would help students invest more time, requiring them to log onto Blackboard more than just once a week, cramming a lesson’s worth of material into a short amount of time (for him, an hour and a half).

Of the four, only Kelsie was putting in the necessary time to actually learn the material and she found she was satisfied with the course. Josh, Ryan and Haley can all be summed up by Haley’s statement, “I think it’s a GPA booster for me” (K. Smith, personal communication, April 20, 2016). Despite not putting in adequate time, Josh, Ryan and Haley all stated they were overall satisfied with the course.

**Technology.** Cole et al. (2014) suggested learning platforms could negatively influence students’ satisfaction in an online class, and although some suggestions for improving online Geol 105 were related to technology, overall, students were positively satisfied with the technology. Kelsie seemed to enjoy the technology the most, from the PowerPoint slides posted online, to the YouTube videos, to the extra credit online scavenger hunt.

One observation that I found particularly interesting, especially as an online instructor myself, is that each of the students used the technology as it best related to their learning style. Kelsie, did not read the textbook but was organized in stepping through the PowerPoint slides
while listening to the audio. As she stated, “If I’m going to do it, I have to be like sitting there 100 percent” (K. Smith, personal communication, April 20, 2016). Haley, on the other hand, described herself as a visual learner and would download the PowerPoint slides, ignore the audio, while taking notes on each slide, “I don’t listen to her recorded lecture, I just click each page and pause it and write the notes. So I copy the PowerPoints basically and that helps me remember it” (H. Adams, personal communication, April 20, 2016). Josh also admitted to ignoring the audio, only reading through the PowerPoint slides without taking notes. He also made a point of stating he did not look at any of the other YouTube videos she posted, “I look at the PowerPoint. I don’t watch her videos” (J. Williams, personal communication, April 20, 2016). His rationale was that, to him, as an elective science, it was not important enough to invest more time than just reading through the PowerPoint, “If it is a class of importance, it absolutely is [worth investing more time]. Like there’s an online video guy who does tutorials, or not tutorials but classes for like the MCAT, DAT and everything. I watch his videos for bio all the time. They’re very helpful. For geology, it’s not something I can pay attention to” (J. Williams, personal communication, April 20, 2016). Ryan described himself as an auditory learner and would often just listen to the PowerPoint audio while doing other work, rarely reading the PowerPoint slides, “I just put [the PowerPoint lecture on] in the background. Like I’ll have like the—MyMathLab or something pulled up while I’m listening to the slides” (R. Jones, personal communication, April 20, 2016). To me, this emphasizes the importance of catering to all learning styles when designing an online class.

One criticism of the use of technology by Kelsie was the requirement of the internet in order to take the weekly quizzes. For her, this was an issue because she typically went home on Fridays and did not have access to the internet. Therefore, she was required to complete her
quizzes before going home each weekend, “I literally live out in the country. It’s just like that’s one inconvenience if like you have to have internet to do [assignments in an online class]” (K. Smith, personal communication, April 20, 2016).

Another criticism regarding technology related to the test scheduling website used for online classes at the university. Geol 105 had two proctored exams, and students were expected to sign up for a time slot on a specific date in order to take their exam at the university’s testing lab. Josh’s recollection of trying to sign up for a time slot that was no longer available was justifiably aggravating, “So one time I was signing up for one [test] late, and I had to go through one 20 times until one of them let me sign up because you have to wait until you select it to know you can sign up” (J. Williams, personal communication, April 20, 2016). He suggested that an easy fix of the test scheduling website was to only show currently available test times, which would save students the frustration of trying to figure out the technology in order to sign up for a proctored test.

Another insightful discussion relating to technology arose when Josh began describing the use of technology in a biology class he was taking. The class used online quizzes in conjunction with an eBook. As he described, if you missed a question on the online quiz, rather than give the answer, Blackboard directed students to the section in the eBook where the topic was discussed. Josh found this to be more confusing than helpful since he found navigating the eBook to be cumbersome. As Josh stated, “It’s terrible. You can’t just have an e-book and go through the pages, you have to like select a chapter and it’ll have a bunch of sections, and you select a section, and it’ll have a bunch of sections, or if you search for an image, a certain figure, it opens it as a PDF, like an outside PDF, that’s not in the book” (J. Williams, personal communication, April 20, 2016). Even though Josh is a Computer Science major who uses
technology extensively in not only the classroom, but also his job on campus, he summarized this courses’ use of technology as “Sometimes people try and use too much technology, and end up ruining things” (J. Williams, personal communication, April 20, 2016).

Self-Regulated Learning Practices. One issue that has been shown to directly relate to student satisfaction are those self-motivating practices students employ in the learning process, and those practices are more pronounced when the learning environment is online (Wang et al., 2013). Of the four students in the focus group, Kelsie and Haley had the best practices in place. For Kelsie, every Monday she would read through the PowerPoint slides and supplementary lecture material. She then followed this routine by taking her mandatory online quiz. She described her strategy for success as “I read it, and read it, when I retake it, some of [the quizzes] I’ll retake four or five times because I’ll get a ninety-five and I want a 100” (K. Smith, personal communication, April 20, 2016). For Haley, she would spend a three hour block of time on Wednesdays while at her job on campus. Ryan said he waited until the last day, Friday, and he often listened to the lecture while trying to do other work. Josh also admitted to doing his work on Friday. I found it interesting that Josh would purposely wait until the last day, but then also criticized the instructor for making assignments due on Friday. Although I did not point this inconsistency out to him, I do not know that any day of the week would have made a difference for him, he still would have pushed taking the quiz to the last possible day.

Convenience. Cole et al. (2014) found convenience to be the predominant reason students enrolled in a course online over face-to-face. In interviewing the Geol 105 students, while convenience was an important issue, it was not as predominant as interaction, technology, and self-regulated learning practices. Haley and Josh were specifically wanting to take this
course online, whereas Kelsie and Ryan only chose the online environment because the face-to-face section was full.

For Haley and Josh, both had heavy course loads for the semester, and they each wanted the flexibility of doing the course on their own time. Both had jobs that allowed them to complete schoolwork during slow times, and often both would try to squeeze in the time to complete the online Geol 105 assignments while working. Another interesting observation is that only Haley and Josh had taken online courses before. Because of their past experiences, they were confident in enrolling in another online course. Kelsie and Ryan opted to take Geol 105 online because the face-to-face course was already filled and they still wanted to take the course in the Spring 2016 semester. Ryan’s attitude was, “I’ve had friends who’ve done online classes, and it seemed like theirs is almost an easier setup than like an in-class part” (R. Jones, personal communication, April 20, 2016). Even though Kelsie and Ryan had not taken a previous course, they did state that after their experience in Geol 105, they would take another online course again, in part, because they liked the flexibility that online courses offer.

Course Structure. Layout of a course was important for all four. Although Kelsie had not taken an online course before, she was quite animated when discussing how impressed she was with the organization of this course, even stating it would help if face-to-face courses were structured similarly, “I wish my in-classes would also post on Blackboard and say here’s your slides. I wish they would have that organization, because then whatever notes I didn’t take in class I can just go look back at it” (K. Smith, personal communication, April 20, 2016). The other three focused more on the weekly quizzes, which could be taken as many times as a student wanted throughout the week with the last quiz submission as the recorded grade. All found the quizzes to be a positive structure to the course, but most admitted this was only true because if
they took the quiz enough times, the questions repeated and you could be guaranteed a perfect score. Moreover, these were similar to the questions on the test, and Kelsie admitted she had seen every question on the class’s proctored test because they were questions from her online quizzes, “I probably see every question, I’m not going to lie” (K. Smith, personal communication, April 20, 2016).

In terms of negative structure to the course, as already mentioned, Haley wished for more opportunities for visual interaction with the teacher. Josh, claimed that he preferred a course structured such that assignments were not always due on Friday: “I was going to say, the whole due Friday, due Friday is very hard for me” (J. Williams, personal communication, April 20, 2016). However, he also admitted to procrastinating to the last minute to complete his quizzes, and the week before the focus group, he actually missed that week’s quiz. Ryan’s suggestion to offer midweek checkpoints as a way to improve the class could possibly be implemented for someone such as Josh to keep him on track to successfully complete weekly assignments, regardless of the day of the week the assignments are due.

Summary

In sum, student satisfaction in the Geol 105 online class for which I conducted my focus group confirms the findings of Cole et al.’s (2014) three year study on student satisfaction in online courses as well as other studies (Armstrong, 2011; Carter & Emerson, 2012, Kuo et al., 2013; Lee, 2014). While the ordering of importance in my study is different than Cole et al’s (2014) study, all themes emerged in the discussion with the four students who participated: interaction, convenience, structure, technology, and self-regulated learning practices. I found the most compelling observation to be that of the students’ impressions of interaction. While they chose to take a class using a remote technological environment, they all felt interaction with the
instructor and other students would enhance the course. This is important information for online instructors to consider when designing such a course, possibly implementing some of the suggestions these students offered as ways of improving students’ experiences in such an environment.
CHAPTER VI

DISCUSSION OF QUANTITATIVE AND QUALITATIVE RESULTS

The chapter begins with a discussion and interpretation of the quantitative results, followed by discussion of the findings of the qualitative results. Finally, discussion concludes with how the qualitative findings broaden the understanding of the quantitative results, and suggestions for future work.

Quantitative Results

**Hypothesis One, Hypothesis Two and Hypothesis Three Conclusions.** For hypothesis one, which postulated performance of the first Geol 105 test, all students, except for exceptionally performing students with high GPAs above 3.91, did significantly better on test one in the face-to-face Geol 105 than in the online Geol 105 after adjusting for GPA. Therefore, the null hypothesis is rejected for students with GPAs below 3.91, whereas there is failure to reject the null hypothesis for students with GPAs 3.91 or higher.

For hypothesis two, the second Geol 105 test, students with middle range GPAs, 2.65 to 3.10, did significantly better on test two in the face-to-face Geol 105 than in the online Geol 105 after adjusting for GPA. Therefore, the null hypothesis is rejected for students with GPAs between 2.65 and 3.10, whereas there is failure to reject the null hypothesis for students with GPAs below 1.96 and above 3.91.

For hypothesis three, the final overall score in Geol 105, all students, except for exceptionally performing students with high GPAs, above 3.92, did significantly better overall in the face-to-face Geol 105 than in the online Geol 105 after adjusting for GPA – this is true for all
years except 2014. Therefore, the null hypothesis is rejected for students with GPAs below 3.92, whereas there is failure to reject the null hypothesis for students with GPAs 3.92 or higher.

**Hypothesis Four Conclusion.** As a result of the regression analysis, the null hypothesis is rejected because for all instances, the complete data and the data by semester, GPA, course load, and completed student credit hours were shown to be significant predictors of final score in the online Geol 105 class. Although hypothesis four had six predictor variables – GPA, prior online experience, major as categorized as STEM and non-STEM, major as categorized as Geology and non-Geology, number of completed student credit hours, and current course load – only GPA, course load and completed student credit hours were shown to be significant.

**Using the OLS Regression Equation in Consideration of the Final Scores ANCOVA.**

For hypothesis three, a robust ANCOVA using R determined that the following GPAs in both course environments, online and face-to-face, had similar relationships between GPA and final score in Geol 105: 1.49, 2.51, 3.01, 3.42, 4.00 (Table 4). In all five cases, the difference in the mean final score was significant and students enrolled in the face-to-face Geol 105 course outperformed those in the online Geol 105 course. An OLS regression was run in SPSS for the face-to-face data, and all regression assumptions were verified.

Table 12

*Face-to-Face SPSS OLS Regression Results*

<table>
<thead>
<tr>
<th>Model 1</th>
<th>B, 95% CI</th>
<th>SE B</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>69.198</td>
<td>0.806</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[67.616, 70.780]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPA</td>
<td>5.925</td>
<td>0.265</td>
<td>.532</td>
</tr>
<tr>
<td></td>
<td>[5.405, 6.445]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The results showed that models 1, 2, and 5 were significant. Because the only model that was common to both course environments was model 1, Table 12 only lists the face-to-face SPSS results for model 1, and the OLS regression equation for the face-to-face Geol 105 course is as follows (p < .001):

\[ F2F \ Geol \ 105 \ final \ score_i = 69.198 + (5.925 \times GPA_i) \]

In evaluating the two OLS regression equations for model 1 (GPA as the only predictor variable), one for the online environment and one for the face-to-face environment, Table 13 shows the calculations for predicting the final score in Geol 105 at each of the five GPAs determined by the ANCOVA for hypothesis three. The results show that students with identical GPAs can expect to do better in the face-to-face Geol 105 course, but the differential between final scores gets smaller as GPA increases. These results confirm results of Jaggars et al.’s (2013b) research in which male students with lower GPAs tended to perform even worse when the course was taken in an online environment over a traditional face-to-face environment.

Table 13

<table>
<thead>
<tr>
<th>GPA</th>
<th>Predicted Geol 105 Final Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Online</td>
</tr>
<tr>
<td>1.49</td>
<td>55.4%</td>
</tr>
<tr>
<td>2.51</td>
<td>70.6%</td>
</tr>
<tr>
<td>3.01</td>
<td>78.1%</td>
</tr>
<tr>
<td>3.43</td>
<td>84.3%</td>
</tr>
<tr>
<td>4.00</td>
<td>92.8%</td>
</tr>
</tbody>
</table>
Summary of the Quantitative Results. Students enrolled in the online Geol 105 class did not perform as well as those students in the face-to-face online Geol 105 class, after controlling for the effect of a student’s incoming GPA. This was shown not only in hypotheses one, two, and three, each of which were evaluated using ANCOVA for the two exams and overall final performance in the class, but also in comparing OLS equations for predicting performance in the online Geol 105 and face-to-face Geol 105 classes.

Why lower performing students, those with low GPAs, perform worse in the online environment can, in part, be explained by the framework of this research study – Moore’s (1993) theory of transactional distance and the self-regulated learning model (Zimmerman & Martinez-Pons, 1986). The adverse impact of distance in the online classroom must be offset by positive interactions, most notably the student-to-content interaction. A student must frequently interact with the material, which requires a student to have self-motivation to regularly log into the course and engage in the content in order to promote learning. The findings from this study suggest that lower performing students struggle to adequately engage in the content and consequently have significantly lower achievement, as much as two letter grades lower, than if these students had taken the course face-to-face instead. For the poorer performing students, the accountability of a live lecture two to three days a week clearly makes a difference in their overall performance.

Higher performing students, however, are minimally impacted, if at all, by the different environments. These students have sufficient self-motivational practices such that they positively interact with the course content and achieve learning regardless of whether the course is taken online or face-to-face. In sum, the quantitative results of this study strongly suggest that a student considering whether to take an online class should first scrutinize his or her
motivational practices to engage in learning, and the first step in an honest examination should be considering one’s past experience in the classroom, most indicative by one’s GPA.

**Qualitative Findings**

Of the four students who participated in the focus group, all described their experiences and satisfaction levels in online Geol 105 in terms of the same findings as Cole et al. (2014): interaction, technology, self-regulated learning practices, convenience, and course structure. While the stronger themes in this study are not the same as Cole et al. (2014), elements of all emerged in the discussion.

**Summary of Qualitative Findings.** Moore’s theory postulating the importance of overcoming the negative impact on learning due to a distance-based classroom, particularly in the interactions students have with the instructor, other students, and the content, was undoubtedly the strongest theme that emerged from the discussion (Moore, 1993). The online instructional designer, one of the coders of the data, noted the importance of students having interaction is summative in Josh’s comment, “I didn’t know her name until now”. As she commented, in spite of busy schedules, students want a connection with the instructor, other students, and the content, and this happens most effectively when there are personal and timely interactions. Clearly, the challenge for online instructors, in order to boost student satisfaction, is to incorporate some mechanism that allows for frequent interactions, possibly implementing some of the suggestions these students offered as ways of improving students’ experiences in such an environment. Findings from this focus group may help to understand why students in the online class did not achieve at the same level as the face-to-face class.

As in previous research, technology can sometimes leave students with a negative impression if they find it difficult to navigate the course (Armstrong, 2011; Cole et al., 2014). In
the situation Josh described in his biology class, he found technology to adversely impact his learning because he felt technology was overused. This suggests that while online course development should implement technology, instructors should carefully consider which technologies to use (Chickering & Ehrmann, 1996). It does not always follow that utilizing educational technology in a course is a pedagogically sound approach.

Regarding technology accessibility, Kelsie expressed a desire that the course allow for offline work. While this criticism is understandably frustrating, because the nature of the course is web-based and it is important to assess student progress in online courses on a consistent basis (for Geol 105, this is done through online weekly quizzes), I do not believe there is a solution to eliminating the requirement of internet access at some point each week, much like courses require a textbook and expect students to have access to it throughout the semester.

Self-regulated learning strategies, no doubt, are important in online classes, but students still want some built-in accountability to encourage them to frequently access the content, much like they would attending lecture each week in a traditional face-to-face class. Several times, Ryan mentioned midweek checkpoints, though he was quick to point out that he did not mean in the form of online quizzes. As an online instructor, I had not considered this, which could be adding something as simple as requiring students to post a comment to an online discussion board midweek.

Convenience was the most broad-sweeping theme, and the other two coders made a point of embellishing on this theme. The online instructional designer observed that two students, Haley and Josh, purposely chose to take the online Geol 105 course, whereas Ryan and Kelsie had different reasons. Ryan stated the other Geol 105 classes were filled, not to mention he did not want to have to take the required lab section for those students enrolling in the face-to-face
Geol 105 course. Kelsie did not necessarily mind the lab, but felt the face-to-face Geol 105 teacher was a poor teacher. The second coder, who also transcribed the focus group, noted that students categorized convenience in terms of other class conflicts, reducing their face-to-face workload, investing in the time required for the class, and believing they could be successful because of previous success in online classes. Clearly, what motivates students to take an online course, as shown by this research, stems from a variety of reasons – some more noteworthy than others. This finding, to me, provides the strongest explanation for the different student performances identified in the quantitative results. Students want the convenience of an online class but do not always consider the discipline required to be academically successful in a web-based course. This reiterates past literature that has suggested the key factors to satisfaction in an online environment are related to self-regulated learning, technology self-efficacy, autonomy, and interaction (Artino & McCoach, 2008; Bolliger & Martindale, 2004; Reinhart & Schneider, 2001).

For the last theme, course structure, the four students enrolled in Geol 105 all approached the convenience and flexibility of the course through the lens of their own learning styles which reinforces Chickering and Gamson’s (1987) last tenet of their seven tenets of good practices in undergraduate education. For Kelsie and Ryan, they consistently listened to the audio in the weekly PowerPoints, whereas Josh and Haley only read the content written in the slides. The online instructional designer noted that this would have been an interesting theme to further explore, including a discussion of their thoughts on how the course structure was designed for those students with hearing and vision-impairment disabilities. In order to promote deep learning and satisfaction in this web delivery mode, careful consideration should be given to organization of online courses as well as to those formats in which lessons are deployed.
Future Work

Understanding the impact of the distance in online classes is of important pedagogical consideration. As shown in previous research, including my own, the debate continues whether one platform is better than the other (Atchley et al., 2013; Carter & Emerson, 2012; Driscoll et al., 2012; Edmonds, 2006; Frantzen, 2013; Hachey et al., 2014; Jaggars et al., 2013b). While this research study used five years of data from just one course, Geol 105, one limitation discussed earlier was that the instructors in each of the environments was different, even though they used the same lecture material and same test bank. Future work should consider data in which both environments are taught by the same instructor over several years. It is important to consider more than just one semester in order to see possible emerging trends in the data. Moreover, if the same instructor is teaching the course, more insight could be given into explaining why differences may have occurred between semesters and/or between environments.

Also, future work conducting more focus groups related to those students enrolled in online classes would help to explain how students describe their satisfaction in online classes, giving light to what may impact student performance. My study had just an 8% response rate, and my hope of garnering volunteers by offering an incentive clearly did not have the success rate that I was hoped it would. One study that had a high volunteer rate was conducted by the instructor of the course, and he offered extra credit for those who participated (Cole et al., 2014). Because I was not the instructor of record, I could not offer the incentive of extra credit, which would likely have improved the 8% response rate. Perhaps if online instructors were willing to participate in educational research related to the online environment, they would consider offering extra credit. Moreover, the purpose of focus groups, in general, is to continue to conduct them until the responses of the participants are predictable, drawing from previous focus
group findings (Vaughn et al., 1996). While I believe my findings complement Cole et al.’s (2014) research, similar focus groups would strengthen these findings.

Finally, future research conducted on other subject areas offered in the online environment, in addition to Geol 105, would contribute to a better understanding of the online delivery mode in higher education. In the focus group, Haley mentioned that she took a course in her major that was only offered online. I, too, have taken such a course. Understanding not only how students perform, but also how they describe their satisfaction in the online environment for such courses would contribute to research in online education.

Conclusion

In regard to previous studies assessing student performance in online classes compared to face-to-face, few have considered predictor variables of student characteristics at the time the course was taken (Carter & Emerson, 2012; Driscoll et al., 2012; Frantzen, 2013). Of these studies, none evaluated repeated semesters of data from the same class that uses the same lecture material and test banks, controlling for students’ actual incoming GPAs, as in my study using five years of Geol 105 data. Although GPA has been shown to have strong significant influence on a student’s performance, regardless of course environment, only three studies controlled for this effect, and they used self-reported GPA rather than actual GPA (Driscoll et al., 2012; Frantzen, 2013; Hachey et al., 2014). My study not only utilized several years of data from the same class with a large sample size, but it also used a student’s actual GPA at the time they took the course (hypotheses one, two and three).

Additionally, predictor variables used in several past studies for OLS regression were from dissimilar online courses, rather than the same course (Atchley et al., 2013; Hachey et al., 2012, Hachey et al., 2014). In my study, evaluating hypothesis four, data was used from the
same course, including predictor variables at the time the students were enrolled in Geol 105: actual GPA, prior online experience, major as categorized as STEM and non-STEM, major as categorized as Geology and non-Geology, number of completed student credit hours, and current course load.

Furthermore, this study has helped to contribute to the ongoing debate of the effectiveness of online education. Many studies have determined there is no difference in student performance by course environment (Carter & Emerson, 2012; Detwiler, 2008; Driscoll et al., 2012; Frantzen, 2013; Summers et al., 2005). Other studies, however, have shown that student performance in the online environment lags that of the face-to-face environment (Edmonds, 2006; Atchley et al., 2013). Findings from this study showed that for students who are already high achieving, those with high incoming GPAs, environment does not impact student performance. On the other hand, for students who are in the middle in terms of GPA, course environment does impact performance, although this difference may not result in a lower letter grade, just a lower overall score. Low-achieving students, however, are decidedly impacted by environment. Students with low GPAs were shown to perform as much as two letter grades lower than those in the face-to-face course.

In terms of how students describe their satisfaction, previous studies have typically administered surveys rather than conducted interviews or focus groups (Armstrong, 2011; Carter & Emerson, 2012; Cole et al., 2014; Kuo et al., 2013; Lee, 2014; Simpson & Benson, 2013; Summer, 2005). Although these past studies have identified prominent themes related to student satisfaction – self-regulated learning, technology, autonomy, and interaction – it is through further discussions on these topics that can help reveal strategies that directly impact how online courses are designed. This study has added to a better understanding of student satisfaction
through the findings from the focus group that explored what students believe impact their level of satisfaction and what they feel would improve this level.

Students choosing to enroll in an online section of a class over the traditional face-to-face class should first examine whether their motivation and self-regulated learning behaviors are conducive to a remote environment. While students seem to be satisfied overall with online courses, it is more important that students choose a course delivery mode by considering their own study habits and established routines rather than by the convenience and flexibility of how lectures and assignments are deployed. The findings from the focus group strongly suggest that students will approach learning in an online class through the lens of their learning styles rather than considering all the material the instructor has posted on the learning platform. Moreover, results from the quantitative data suggest that a student’s past history with college courses is by far the strongest predictor of how they can expect to perform, and this is more pronounced in the online environment for lower GPAs. As educational technologies continue to evolve, especially relating to the online classroom, it is critical that educators provide information to students such that they can make the best decision in choosing the best learning platform for their own academic success.
LIST OF REFERENCES


online versus a face-to-face introductory Sociology course. *Teaching Sociology, 40*(4), 312-331.


Kranzow, J. (2013). Faculty leadership in online education: Structuring courses to impact student satisfaction and persistence. *MERLOT Journal of Online Learning and Teaching, 9*(1), 131-139.


Lee, J. (2014). An exploratory study of effective online learning: Assessing satisfaction of levels of graduate students of mathematics education associated with human and design factors of an online course. *International Review of Research in Open and Distance Learning, 15*(1), 111-132.


Pence, H. E. (2012). When will college truly leave the building: If MOOCs are the answer, what is the question? *Journal of Educational Technology Systems, 41*(1), 25-33.


LIST OF APPENDICES
APPENDIX A: EMAIL CORRESPONDENCE
IRB Approval

Ms. Davidson:

This is to inform you that your application to conduct research with human participants, “Comparing student performance and investigating student perceptions of satisfaction in university online course: Does course environment make a difference? " (Protocol #16x-201), has been approved as Exempt under 45 CFR 46.101(b)(#2, 4).

Please remember that all of The University of Mississippi’s human participant research activities, regardless of whether the research is subject to federal regulations, must be guided by the ethical principles in The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research.

It is especially important for you to keep these points in mind:

• You must protect the rights and welfare of human research participants.

• Any changes to your approved protocol must be reviewed and approved before initiating those changes.

• You must report promptly to the IRB any injuries or other unanticipated problems involving risks to participants or others.

If you have any questions, please feel free to contact the IRB at irb@olemiss.edu.

Jennifer Caldwell, PhD
Senior Research Compliance Specialist, Research Integrity and Compliance
The University of Mississippi

Clarification of IRB Approval

Please note that you will not need to obtain signatures for the focus groups, as this is an exempt study.

Sending approval separately.

Jennifer Caldwell, PhD
Senior Research Compliance Specialist, Research Integrity and Compliance

Email to Instructors to Update My Research Status

Just an update for you that I did pass my PhD prospectus yesterday. The next step is for me to file an IRB request to use the data from Geol 105. I've been told that it takes about
Email to Instructors Requesting Their Data

I did receive IRB approval for my PhD research:
This is to inform you that your application to conduct research with human participants, “Comparing student performance and investigating student perceptions of satisfaction in university online course: Does course environment make a difference? ” (Protocol #16x-201), has been approved as Exempt under 45 CFR 46.101(b)(2, 4).

What I need in the Box folders that your Chair has created for each of you is just the student ID's along with test grades, final exam grade, final grade, and final letter grade. If you would, name the files according to the semester in which they were given. Many thanks,
Kristi

Email to Institutional Research

I have received IRB approval for my research study:
This is to inform you that your application to conduct research with human participants, “Comparing student performance and investigating student perceptions of satisfaction in university online course: Does course environment make a difference? ” (Protocol #16x-201), has been approved as Exempt under 45 CFR 46.101(b)(2, 4).
I am now ready for Institutional Research's assistance. Since we've met once already, I'll outline the details in this email. If you think we should meet again, just let me know and we can schedule a time - I'm free Tuesdays and Thursdays and afternoons most Mondays, Wednesdays, and Fridays.

You will be receiving an email from Geological Engineering that you have been added to 2 folders on Box. One is for the online Geology 105, and the other for the traditional face-to-face Geology 105. The online instructor told me she included all her semesters, including summers, and put them in 1 Excel Workbook with tabs delineating the semesters. The face-to-face told me she just pulled the information directly off of Blackboard, so, for her, each semester is in its own file. When you append data, if you would, keep them by semesters. That way when I do the propensity score matching, I can do it by semester.

So, what we had discussed and what my IRB gives me exemption for is the following:
Beginning of the Semester Stats

All of this information is at the time they enrolled to take Geol 105. For some, it may have been their last semester, but for most it will not be.

1. GPA
2. Number of previously taken online courses and grades associated with those courses
3. Declared Major
4. Number of completed student credit hours (or course classification)

Ending Semester Stats

1. Course load for the semester
2. Letter grade in the course

Demographics

1. Gender
2. Ethnicity
3. Age

Again, please let me know if you have any questions.

Thanks,
Kristi

Email to Online Instructor Regarding Focus Group

After spring break, I will be sending you the email I'd like you to send out to your online students, requesting volunteers for my focus group (I will be giving $25 Amazon gift cards to all of the participants - incentive for getting volunteers). I am hoping to get 8-12 students, and, even better, would be a group of 8-12 representing a diversity in performance.

Right now, I'm just in the waiting game. Institutional research is appending to the Excel files student GPA, major, number of online courses (and grades in each), etc. That part will give me the statistics. The focus group interview is to help explain the statistics. Kristi

Email to Solicit Student Volunteers for Focus Group

All:

Would you like to earn a $25 Amazon gift card for 2 hours of your time by helping with a research study?

I am a doctoral student in Higher Education here at Ole Miss. Your instructor, Ms. Patterson has agreed to help me out!
I am currently researching student satisfaction in online classes and am looking for volunteers to discuss this in a small group. What you need to know:

* This is strictly voluntary. It in no way will help or harm your grade in the course if you choose to participate or not. In fact, Ms. Patterson will have no idea who is in the group and who is not.

* The discussion time will last about 2 hours on the main campus.

* If you do participate, you will receive a $25 Amazon gift card for your time.

* The group will consist of about 8-12 students, all of you in the online Geol 105 class this semester. If more students than I need are interested in the focus group, please keep in mind I will select those 8-12 students that will create the most diverse group.

* Questions and discussion topics will be related to what you have liked and disliked regarding the online environment.

If you are interested or have any questions, please email me at kdavidso@olemiss.edu

I will be putting together a list of names and will get back to you soon on when and where the group discussion will be conducted.

Thanks so much for your consideration!

Kristi Davidson, M.S.
Doctoral Student
Department of Higher Education

Email to Student Volunteers that Responded to the Email

Fantastic!! I will be in touch in the next couple of weeks (or sooner if I get enough responses) about when and where we will meet.

Thanks so much for agreeing to participate!
Kristi

Email to Determine a Time and Date for the Focus Group

All:
I wanted to thank you again for agreeing to participate in the focus group for my PhD research on your experience in taking an online class. Once we have decided on a date and time, I will send you some types of questions I’d like to discuss (this is in addition to
any additional comments you may have about online classes). As I mentioned in my original email, I do not plan to take more than 2 hours of your time, and the actual focus group discussion I expect to take no longer than 1 - 1.5 hours. When you come for the discussion, I’d like to first get some information from you such as what year you are in school, what your major is, etc. I will also have your Amazon gift card at that time.

My preference is to conduct the group interview next week, Tuesday, April 19. However, rather than set a day and time, I’d like you to fill in the times below in which you are available to meet (keep in mind you need a 2 hour block of time).

Once I hear from everyone, I’ll reply back with a time. Also, let me know if evenings work better, and what times. If it happens that schedules conflict, we may have to do an evening time, though this would be a last resort.

Thanks,
Kristi Davidson

Monday, April 18:

Tuesday, April 19:

Wednesday, April 20:

Thursday, April 21:

Friday, April 22:

Email to Participants with Time and Location of Focus Group

Again, many thanks for agreeing to participate in the focus group. I have heard back from everyone and a time has been set! It will be next week, Wednesday, April 20 from 3 - 4:30 in Weir Hall 225 (second floor of the same building as the help desk). Below you will find a list of questions that I would like to discuss, but by all means we are not limited to these in our discussion. I will see you all next week and please do not hesitate to contact me with any questions or concerns.

Thanks,
Kristi

Possible Questions to Discuss:
1. Why did you enroll in the online version of Geol 105 rather than the traditional, face-to-face version?

2. How many online classes have you taken before?

3. How does Geol 105 compare to these other courses?

4. When you reflect on your learning style, do you think online was a good choice, or would the traditional face-to-face have been better? Why?

5. Did you find your instructor provided prompt and helpful responses to your questions?

6. How comfortable were you using Blackboard?

7. Was any other technology such as YouTube or another website used besides Blackboard?

8. What would you identify as the positive characteristics of having taken this course online?

9. What would you identify as the negative characteristics of having taken this course online?

10. Would you take another online class? Why or why not?

11. Overall were you satisfied with Geol 105?

12. What would have made Geol 105 even better?
APPENDIX B: FOCUS GROUP SEMI-STRUCTURED INTERVIEW QUESTIONS
Focus Group Questions

1. Why did you enroll in the online version of Geol 105 rather than the traditional, face-to-face version?

2. How many online classes have you taken before?

3. How does Geol 105 compare to these other courses?

4. When you reflect on your learning style, do you think online was a good choice, or would the traditional face-to-face have been better? Why?

5. Did you find your instructor provided prompt and helpful responses to your questions?

6. How comfortable were you using Blackboard?

7. Was any other technology such as YouTube or another website used besides Blackboard?

8. What would you identify as the positive characteristics of having taken this course online?

9. What would you identify as the negative characteristics of having taken this course online?

10. Would you take another online class? Why or why not?

11. Overall were you satisfied with Geol 105?

12. What would have made Geol 105 even better?
APPENDIX C: R OUTPUT FOR ROBUST ANCOVA AND ROBUST BOOTSTRAP ANCOVA
Robust ANCOVA R Results for Test 1 Using All Data

```R
> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
[1] "NOTE: Confidence intervals are adjusted to control the probability"
[1] "of at least one Type I error."
[1] "But p-values are not"

\[
\begin{array}{cccc}
\text{X} & \text{n1} & \text{n2} & \text{DIF} & \text{TEST} \\
\hline
1 & 1.49 & 18 & 186 & -13.742351 & 3.971237 \\
2 & 2.51 & 105 & 722 & -16.900870 & 10.907340 \\
3 & 3.01 & 132 & 900 & -14.480420 & 10.743443 \\
4 & 3.43 & 108 & 836 & -10.752934 & 7.528456 \\
5 & 4.00 & 57 & 489 & -6.248988 & 4.672061 \\
\end{array}
\]

\[
\begin{array}{ccc}
\text{se} & \text{ci.low} & \text{ci.hi} \\
\hline
1 & 3.460471 & -24.102892 \\
2 & 1.549495 & -20.984778 \\
3 & 1.347838 & -18.006914 \\
4 & 1.428305 & -14.518469 \\
5 & 1.337523 & -9.864142 \\
\end{array}
\]

\[
\begin{array}{c}
p\text{.value} & \text{crit.val} \\
\hline
1 & 1.742348e-03 & 2.993968 \\
2 & 0.000000e+00 & 2.635638 \\
3 & 0.000000e+00 & 2.616409 \\
4 & 1.556764e-10 & 2.636365 \\
5 & 4.047306e-05 & 2.702872 \\
\end{array}
\]

> ancboot(covGrp1, dvGrp1, covGrp2, dvGrp2, nboot=2000)
[1] "Note: confidence intervals are adjusted to control FWE" 
[1] "But p-values are not adjusted to control FWE" 

\[
\begin{array}{cccc}
\text{X} & \text{n1} & \text{n2} & \text{DIF} & \text{TEST} \\
\hline
1 & 1.49 & 18 & 186 & -13.742351 & 3.971237 \\
2 & 2.51 & 105 & 722 & -16.900870 & 10.907340 \\
3 & 3.01 & 132 & 900 & -14.480420 & 10.743443 \\
4 & 3.43 & 108 & 836 & -10.752934 & 7.528456 \\
5 & 4.00 & 57 & 489 & -6.248988 & 4.672061 \\
\end{array}
\]

\[
\begin{array}{ccc}
\text{ci.low} & \text{ci.hi} & p\text{.value} \\
\hline
1 & -23.286080 & -4.198622 \\
2 & -21.174266 & -12.627475 \\
3 & -18.197659 & -10.763180 \\
4 & -14.692098 & -6.813771 \\
5 & -9.937779 & -2.560196 \\
\end{array}
\]

\[
\begin{array}{c}
\text{crit} \\
\hline
1 & 2.757928 \\
\end{array}
\]

> 
```
Robust ANCOVA R Results for Test 1 for 2011

```R
> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
```

```
[1] "NOTE: Confidence intervals are adjusted to
control the probability"
[1] "of at least one Type I error."
[1] "But p-values are not"
```

<table>
<thead>
<tr>
<th>X</th>
<th>n1</th>
<th>n2</th>
<th>DIF</th>
<th>TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.96</td>
<td>19</td>
<td>70</td>
<td>-13.287582</td>
</tr>
<tr>
<td>2</td>
<td>2.36</td>
<td>25</td>
<td>101</td>
<td>-11.318623</td>
</tr>
<tr>
<td>3</td>
<td>2.65</td>
<td>29</td>
<td>113</td>
<td>-12.452891</td>
</tr>
<tr>
<td>4</td>
<td>3.14</td>
<td>24</td>
<td>102</td>
<td>-9.403286</td>
</tr>
<tr>
<td>5</td>
<td>3.78</td>
<td>13</td>
<td>60</td>
<td>-9.024444</td>
</tr>
</tbody>
</table>

```

<table>
<thead>
<tr>
<th>se</th>
<th>ci.low</th>
<th>ci.hi</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.399956</td>
<td>-23.08426</td>
</tr>
<tr>
<td>2</td>
<td>3.065636</td>
<td>-20.01237</td>
</tr>
<tr>
<td>3</td>
<td>2.683842</td>
<td>-19.85901</td>
</tr>
<tr>
<td>4</td>
<td>2.262457</td>
<td>-15.71765</td>
</tr>
<tr>
<td>5</td>
<td>1.679956</td>
<td>-14.09177</td>
</tr>
</tbody>
</table>

```

<table>
<thead>
<tr>
<th>p.value</th>
<th>crit.val</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.202269e-03</td>
</tr>
<tr>
<td>2</td>
<td>3.59788e-03</td>
</tr>
<tr>
<td>3</td>
<td>9.207395e-05</td>
</tr>
<tr>
<td>4</td>
<td>4.072483e-04</td>
</tr>
<tr>
<td>5</td>
<td>1.760417e-04</td>
</tr>
</tbody>
</table>
```

```R
> ancboot(covGrp1, dvGrp1, covGrp2, dvGrp2, nboot=2000)
```

```
[1] "Note: confidence intervals are adjusted to
control FWE"
[1] "But p-values are not adjusted to control FWE"
```

```
<table>
<thead>
<tr>
<th>X</th>
<th>n1</th>
<th>n2</th>
<th>DIF</th>
<th>TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.96</td>
<td>19</td>
<td>70</td>
<td>-13.287582</td>
</tr>
<tr>
<td>2</td>
<td>2.36</td>
<td>25</td>
<td>101</td>
<td>-11.318623</td>
</tr>
<tr>
<td>3</td>
<td>2.65</td>
<td>29</td>
<td>113</td>
<td>-12.452891</td>
</tr>
</tbody>
</table>
```

136
Robust ANCOVA R Results for Test 1 for 2012

```r
> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
```

```
[1] "NOTE: Confidence intervals are adjusted to control the probability"
```

```
[1] "of at least one Type I error."
```

```
[1] "But p-values are not"
```

```
$crit
[1] 2.858301
```

```r
> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
```

```
x n1 n2 DIF TEST
[1,] 1.92 14 92 -20.06437 3.920217
[2,] 2.48 30 122 -17.33718 5.895552
[3,] 2.90 38 135 -17.50926 6.730848
[4,] 3.35 30 114 -13.99278 4.410361
[5,] 3.92 14 63 -10.53269 3.696129

se ci.low ci.hi
[1,] 5.118180 -36.04986 -4.078889
[2,] 2.940722 -25.63288 -9.041470
[3,] 2.601345 -24.64504 -10.373483
[4,] 3.172705 -22.97155 -5.014007
[5,] 2.849655 -19.09398 -1.971408

p.value crit.val
[1,] 3.071595e-03 3.123275
```
Robust ANCOVA R Results for Test 1 for 2014

> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)

[1] "NOTE: Confidence intervals are adjusted to control the probability"
"of at least one Type I error."

"But p-values are not"

Note: confidence intervals are adjusted to control FWE

Taking bootstrap samples. Please wait.

> ancboot(covGrp1, dvGrp1, covGrp2, dvGrp2, nboot=2000)

$crit

[1] 3.738184
Robust ANCOVA R Results for Test 1 for 2015

```r
> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
[1] "NOTE: Confidence intervals are adjusted to control the probability"
[1] "of at least one Type I error."
[1] "But p-values are not"

```

```
$output
X n1 n2         DIF      TEST
[1,] 2.27 14  85 -18.5535294 4.5876048
[2,] 2.80 23 128 -15.5085256 5.1477484
[3,] 3.33 31 148 -6.0500234 2.3212726
[4,] 3.57 26 134 -3.2438415 1.4807799
[5,] 4.00 18  93   0.2077193 0.1313129
    se     ci.low     ci.hi
[1,] 4.044274 -31.018080 -6.088979
[2,] 3.012681 -24.195811 -6.821241
[3,] 2.606339 -13.395941  1.295895
[4,] 2.190630  -9.508250  3.020567
[5,] 1.581865  -4.386408  4.801846
    p.value  crit.val
[1,] 9.164656e-04  3.082025
[2,] 9.194546e-05  2.883572
[3,] 3.089278e-02  2.818481
[4,] 1.564716e-01  2.859637
[5,] 8.972308e-01  2.904247
```

```r
> ancboot(covGrp1, dvGrp1, covGrp2, dvGrp2, nboot=2000)
[1] "Note: confidence intervals are adjusted to control FWE"
[1] "But p-values are not adjusted to control FWE"

```

```
$output
X n1 n2         DIF      TEST
[1,] 2.27 14  85 -18.5535294 -4.5876048
[2,] 2.80 23 128 -15.5085256 -5.1477484
[3,] 3.33 31 148 -6.0500234 -2.3212726
```

140
Robust ANCOVA R Results for Test 2 Using All Data

> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
[1] "NOTE: Confidence intervals are adjusted to control the probability"
[1] "of at least one Type I error."
[1] "But p-values are not"
$output

X  n1  n2        DIF     TEST
[1,] 1.49  18 186 -11.824613 2.756026
[2,] 2.51 105 722 -10.412304 6.727769
[3,] 3.01 132 900  -9.602710 8.258566
[4,] 3.43 108 836  -7.313141 5.788241
[5,] 4.00  57 489  -4.669293 2.993825

se     ci.low      ci.hi
[1,] 4.290458  -24.704730  1.0555038
[2,] 1.547661  -14.488282  6.3363266
[3,] 1.162758  -12.637150 -6.5682695
[4,] 1.263448  -10.638591 -3.9876903

$crit
[1] 3.216179

Test 2
> ancboot(covGrp1, dvGrp1, covGrp2, dvGrp2, nbo
t=2000)
[1] "Note: confidence intervals are adjusted to
control FWE"
[1] "But p-values are not adjusted to control F
WE"
$output
   X  n1  n2        DIF      TEST
[1,] 1.49  18 186 -11.824613 -2.756026
[2,] 2.51 105 722 -10.412304 -6.727769
[3,] 3.01 132 900 -9.602710  8.258566
[4,] 3.43 108 836 -7.313141 -5.788241
[5,] 4.00  57 489 -4.669293 -2.993825
   ci.low      ci.hi  p.value
[1,] -23.993911  0.3446848  0.028
[2,] -14.802032 -6.0225764  0.000
[3,] -12.900713 -6.3047071  0.000
[4,] -10.896738 -3.7295439  0.000
[5,]  -9.093002 -0.2455835  0.005
$crit
[1] 2.836363

>
Robust ANCOVA R Results for Test 2 for 2011

> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
[1] "NOTE: Confidence intervals are adjusted to control the probability"
[1] "of at least one Type I error."
[1] "But p-values are not"
$output
$X n1 n2        DIF       TEST
[1,] 1.96 19  70 -6.7564103 1.7290821
[2,] 2.36 25 101 -6.7016393 1.8148023
[3,] 2.65 29 113 -6.8672769 2.2095462
[4,] 3.14 24 102 -1.2580645 0.4221767
[5,] 3.78 13  60  0.8611111 0.2887691
se ci.low     ci.hi
[1,] 3.907513 -18.154441  4.641621
[2,] 3.692766 -17.386522  3.983243
[3,] 3.108003 -15.562451  1.827897
[4,] 2.979947  -9.745581  7.229452
[5,] 2.982005  -8.655049 10.377272
$p.value crit.val
[1,] 0.10451190 2.916953
[2,] 0.08853251 2.893463
[3,] 0.03804994 2.797672
[4,] 0.67786613 2.848210
[5,] 0.77948769 3.191195

> ancboot(covGrp1, dvGrp1, covGrp2, dvGrp2, nboot=2000)
[1] "Note: confidence intervals are adjusted to control FWE"
[1] "But p-values are not adjusted to control FWE."
$output
$X n1 n2        DIF       TEST
[1,] 1.96 19  70 -6.7564103 1.7290821
[2,] 2.36 25 101 -6.7016393 1.8148023
[3,] 2.65 29 113 -6.8672769 2.2095462
[4,] 3.14 24 102 -1.2580645 0.4221767
[5,] 3.78 13  60  0.8611111 0.2887691
se ci.low     ci.hi p.value
[1,] 3.514583

\$crit
[1] 3.514583
Robust ANCOVA R Results for Test 2 for 2012

```r
> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
[1] "NOTE: Confidence intervals are adjusted to control the probability"
[1] "of at least one Type I error."
[1] "But p-values are not"

$output
           X n1 n2   DIF     TEST
[1,] 1.92 14  92 -10.362500 1.618135
[2,] 2.48 30 122 -5.389670 1.776719
[3,] 2.90 38 135 -8.061728 3.668438
[4,] 3.35 30 114 -8.744333 3.628074
[5,] 3.92 14  63 -8.145897 3.257564

            se  ci.low  ci.hi
[1,] 6.403978 -30.39222  9.667216
[2,] 3.033495  -3.90140  3.122063
[3,] 2.197592  -14.03559 -2.087864
[4,] 2.410186  -15.45800 -2.030669
[5,] 2.500610  -15.59908 -0.692710

            p.value crit.val
[1,] 0.1380445708  3.127699
[2,] 0.0901149376  2.805916
[3,] 0.0008703045  2.718369
[4,] 0.0014399124  2.785538
[5,] 0.0063704714  2.980547

> ancboot(covGrp1, dvGrp1, covGrp2, dvGrp2, nboot=2000)
[1] "Note: confidence intervals are adjusted to control FWE"
[1] "But p-values are not adjusted to control FWE"

$output
           X n1 n2   DIF     TEST
[1,] 1.92 14  92 -10.362500 -1.618135
[2,] 2.48 30 122 -5.389670  1.776719
[3,] 2.90 38 135 -8.061728  3.668438
```
Robust ANCOVA R Results for Test 2 for 2014

```r
> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
[1] "NOTE: Confidence intervals are adjusted to control the probability"
[1] "of at least one Type I error."
[1] "But p-values are not"

$crit
[1] 3.095913
```

```
Robust ANCOVA R Results for Test 2 for 2014

> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
[1] "NOTE: Confidence intervals are adjusted to control the probability"
[1] "of at least one Type I error."
[1] "But p-values are not"

$crit
[1] 3.095913
```
> ancboot(covGrp1, dvGrp1, covGrp2, dvGrp2, nboot=2000)
[1] "Note: confidence intervals are adjusted to control FWE"
[1] "But p-values are not adjusted to control FWE"
$output
X  n1  n2        DIF      TEST
[1,] 2.15 12 121 -14.551370 -3.998475
[2,] 2.65 27 212 -12.612132 -4.544130
[3,] 3.10 29 257 -11.122581 -4.549512
[4,] 3.36 28 252 -9.255848 -3.911151
[5,] 3.91 12 154 -5.542553 -1.957036
  ci.low  ci.hi p.value
[1,] -27.57381 -1.5289325  0.0065
[2,] -22.54376 -2.6805012  0.0000
[3,] -19.87088 -2.3742839  0.0000
[4,] -17.72411 -0.7875826  0.0025
[5,] -15.67686  4.5917490  0.1020
$crit
[1] 3.57835

>
Robust ANCOVA R Results for Test 2 for 2015

> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
[1] "NOTE: Confidence intervals are adjusted to control the probability"
[1] "of at least one Type I error."
[1] "But p-values are not"
$output

X n1  n2         DIF       TEST
[1,] 2.27 14  85 -10.6117647 2.58456201
[2,] 2.80 23 128 -7.5974359 2.78014451
[3,] 3.33 31 148 -2.2116959 1.16609194
[4,] 3.57 26 134 -0.2317073 0.09719104
[5,] 4.00 18  93  2.2368421 0.94283621

se  ci.low     ci.hi   p.value crit.val
[1,] 4.105827 -23.123016 1.8994861  0.02527787 3.047194
[2,] 2.732749 -15.356905 0.1620333  0.01206740 2.839438
[3,] 1.896674  -7.422566 2.9991740  0.25375182 2.747373
[4,] 2.384040  -6.964480 6.5010655  0.92355666 2.824102
[5,] 2.372461  -4.652520 9.1262039  0.36033172 2.903888

> ancboot(covGrp1, dvGrp1, covGrp2, dvGrp2, nboot=2000)
[1] "Note: confidence intervals are adjusted to control FWE"
[1] "But p-values are not adjusted to control FWE"
$output

X n1  n2         DIF       TEST
[1,] 2.27 14  85 -10.6117647 -2.58456201
[2,] 2.80 23 128 -7.5974359 -2.78014451
[3,] 3.33 31 148 -2.2116959 -1.16609194
[4,] 3.57 26 134 -0.2317073 -0.09719104
[5,] 4.00 18  93  2.2368421  0.94283621

  ci.low     ci.hi   p.value
[1,] -22.965482  1.7419522  0.0305
[2,] -15.819799  0.6249274  0.0130
[3,] -7.918456  3.4950641  0.2505
[4,] -7.404867  6.9414524  0.9280
[5,] -4.901479  9.3751630  0.3505

$crit
[1] 3.008825
Final Score

Robust ANCOVA R Results for Final Score Using All Data

```r
> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
[1] "NOTE: Confidence intervals are adjusted to control the probability"
[1] "of at least one Type I error."
[1] "But p-values are not"
```

```
$output
   X   n1  n2        DIF     TEST
[1,] 1.49  18 186 -11.817607  3.437154
[2,] 2.51 105 722 -10.079444  8.797544
[3,] 3.01 132 900 -8.518641  8.972324
[4,] 3.43 108 836 -6.110729  6.447398
[5,] 4.00  57 489 -3.277441  3.067800
   se  ci.low      ci.hi
[1,] 3.4381951 -22.226424 -1.4087904
[2,] 1.1457111  9.604175  11.034042
[3,] 1.0390381  6.570877 10.999137
[4,] 0.9477822  3.606480  6.606480
[5,] 1.0683361 -0.204909  6.170404
   p.value crit.val
[1,] 5.190634e-03  3.027407
[2,] 6.441514e-13  2.634107
[3,] 4.218847e-14  2.612602
[4,] 1.231894e-08  2.633254
[5,] 4.157749e-03  2.707915
```

```r
> ancboot(covGrp1, dvGrp1, covGrp2, dvGrp2, nboot=2000)
[1] "Note: confidence intervals are adjusted to control FWE"
[1] "But p-values are not adjusted to control FWE"
```

```
$output
   X   n1  n2        DIF     TEST      ci.low      ci.hi  p.value
[1,] 1.49  18 186 -11.817607 -3.437154 -21.388194 -2.2470201  0.0035
[2,] 2.51 105 722 -10.079444 -8.797544 -13.268654 -6.8902335  0.0000
```
Robust ANCOVA R Results for Final Score for 2011

> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
[1] "NOTE: Confidence intervals are adjusted to control the probability"
[1] "of at least one Type I error."
[1] "But p-values are not"
$crit
[1] 2.783608

$output
   X n1 n2  DIF    TEST
[1,] 1.96 19  70 -10.663297  2.888356
[2,] 2.36 25 101 -9.056404  2.638716
[3,] 2.65 29 113 -8.881709  3.019703
[4,] 3.14 24 102 -3.584496  1.322924
[5,] 3.78 13  60 -2.946667  1.576288

   se  ci.low  ci hi
[1,] 3.691823 -21.589254  0.2626606
[2,] 3.432126  -19.056404  0.9147264
[3,] 2.941253  -17.881709  3.0197030
[4,] 2.709525  -11.358449  4.1533278
[5,] 1.869371  -8.699269  2.8059354
> ancboot(covGrp1, dvGrp1, covGrp2, dvGrp2, nbo ot=2000)
[1] "Note: confidence intervals are adjusted to control FWE"
[1] "But p-values are not adjusted to control F
WE"
$output
X n1 n2        DIF      TEST
[1,] 1.96 19  70 -10.663297 -2.888356
[2,] 2.36 25 101 -9.056404 -2.638716
[3,] 2.65 29 113 -8.881709 -3.019703
[4,] 3.14 24 102 -3.584496 -1.322924
[5,] 3.78 13  60 -2.946667 -1.576288
$crit
[1] 2.955267

>
Robust ANCOVA R Results for Final Score for 2012

```r
> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
```

[1] "NOTE: Confidence intervals are adjusted to control the probability"

[1] "of at least one Type I error."

[1] "But p-values are not"

$output

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x1</td>
<td>n1</td>
<td>n2</td>
<td>DIF</td>
<td>TEST</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>--------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.92</td>
<td>14</td>
<td>92</td>
<td>-13.838929</td>
<td>3.610547</td>
</tr>
<tr>
<td>2</td>
<td>2.48</td>
<td>30</td>
<td>122</td>
<td>-9.453003</td>
<td>4.823128</td>
</tr>
<tr>
<td>3</td>
<td>2.90</td>
<td>38</td>
<td>135</td>
<td>-9.546049</td>
<td>5.293224</td>
</tr>
<tr>
<td>4</td>
<td>3.35</td>
<td>30</td>
<td>114</td>
<td>-8.249000</td>
<td>4.378839</td>
</tr>
<tr>
<td>5</td>
<td>3.92</td>
<td>14</td>
<td>63</td>
<td>-7.022308</td>
<td>3.414234</td>
</tr>
</tbody>
</table>

se ci.low     ci.hi
[1,] 3.832917 -25.82615 -1.851709
[2,] 1.959932 -14.96856 -3.937448
[3,] 1.803447 -14.46761 -4.624491
[4,] 1.883833 -13.49957 -2.998431
[5,] 2.056774 -13.22751 -0.817101

p.value crit.val
[1,] 5.103621e-03 3.127440
[2,] 9.819511e-05 2.814156
[3,] 9.990308e-06 2.728973
[4,] 2.95527e-04 2.787174
[5,] 5.243072e-03 3.016961

> ancboot(covGrp1, dvGrp1, covGrp2, dvGrp2, nboot=2000)
```

[1] "Note: confidence intervals are adjusted to control FWE"

[1] "But p-values are not adjusted to control FWE"


$output

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x1</td>
<td>n1</td>
<td>n2</td>
<td>DIF</td>
<td>TEST</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>--------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.92</td>
<td>14</td>
<td>92</td>
<td>-13.838929</td>
<td>3.610547</td>
</tr>
<tr>
<td>2</td>
<td>2.48</td>
<td>30</td>
<td>122</td>
<td>-9.453003</td>
<td>4.823128</td>
</tr>
<tr>
<td>3</td>
<td>2.90</td>
<td>38</td>
<td>135</td>
<td>-9.546049</td>
<td>5.293224</td>
</tr>
<tr>
<td>4</td>
<td>3.35</td>
<td>30</td>
<td>114</td>
<td>-8.249000</td>
<td>4.378839</td>
</tr>
<tr>
<td>5</td>
<td>3.92</td>
<td>14</td>
<td>63</td>
<td>-7.022308</td>
<td>3.414234</td>
</tr>
</tbody>
</table>

ci.low     ci.hi p.value
[1,] -25.95990 -1.7179594 0.0185
[2,] -15.65096 -3.2550414 0.0000
[3,] -15.24915 -3.8429450 0.0000
[4,] -14.20631 -2.2916895 0.0000
[5,] -13.52652 -0.5180983 0.0055

$crit
[1] 3.162335

>
Robust ANCOVA R Results for Final Score for 2014

```r
> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
[1] "NOTE: Confidence intervals are adjusted to control the probability"
[1] "of at least one Type I error."
[1] "But p-values are not"
$output
          X  n1  n2       DIF       TEST
 [1,]  2.15 12 121 -9.948271  1.888147
 [2,]  2.65 27 212 -9.397569  4.229465
 [3,]  3.10 29 257 -7.973056  3.764611
 [4,]  3.36 28 252 -6.725768  3.806052
 [5,]  3.91 12 154 -2.465957  1.271475

     se  ci.low  ci hi
 [1,] 5.268800 -27.640544  7.744003
 [2,] 2.221928 -15.754400  1.689774
 [3,] 2.117896 -13.968428  1.977684
 [4,] 1.767125 -11.761761  1.686974
 [5,] 1.939447 -8.958029  4.026114
	p.value  crit.val
 [1,] 0.0999851191  3.357932
 [2,] 0.0005387528  2.860952
 [3,] 0.0012890734  2.830815
 [4,] 0.0012926728  2.849824
 [5,] 0.2428485836  3.347383

> ancboot(covGrp1, dvGrp1, covGrp2, dvGrp2, nboot=2000)
[1] "Note: Confidence intervals are adjusted to control FWE"
[1] "But p-values are not adjusted to control FWE"
$output
          X  n1  n2       DIF       TEST
 [1,]  2.15 12 121 -9.948271 -1.888147
 [2,]  2.65 27 212 -9.397569 -4.229465
 [3,]  3.10 29 257 -7.973056 -3.764611
```
$crit$
[1] 4.59669

Robust ANCOVA R Results for Final Score for 2015

> ancova(covGrp1, dvGrp1, covGrp2, dvGrp2)
[1] "NOTE: Confidence intervals are adjusted to control the probability"  
[1] "of at least one Type I error."  
[1] "But p-values are not"

$output$
X n1 n2 DIF TEST
[1,] 2.27 14  85 -8.8131961 3.4450619
[2,] 2.80 23 128 -6.7919487 3.8922079
[3,] 3.33 31 148 -2.1456725 1.3259255
[4,] 3.57 26 134 -0.6801677 0.3750934
[5,] 4.00 18  93  2.0049123 1.8652028

se ci.low ci.hi p.value crit.val
[1,] 2.558211 -16.601217 -1.025175
[2,] 1.745012 -11.756973 -1.826925
[3,] 1.618245 -6.671123  2.379778
[4,] 1.813329 -5.864442  4.504107
[5,] 1.074903 -1.189548  5.199373

p.value crit.val
[1,] 0.005384266 3.044323
[2,] 0.001041022 2.845267
> ancboot(covGrp1, dvGrp1, covGrp2, dvGrp2, nbot=2000)

1 "Note: confidence intervals are adjusted to control FWE"
2 "But p-values are not adjusted to control FWE"
3 "Taking bootstrap samples. Please wait."

$output

 X n1 n2        DIF       TEST
[1,] 2.27 14  85 -8.8131961 -3.4450619
[2,] 2.80 23 128 -6.7919487 -3.8922079
[3,] 3.33 31 148 -2.1456725 -1.3259255
[4,] 3.57 26 134 -0.6801677  0.3750934
[5,] 4.00 18  93  2.0049123  1.8652028

$ci.low     $ci.hi  $p.value
[1,] -16.594648 -1.031745  0.0025
[2,] -12.099847 -1.484051  0.0025
[3,]  -7.067978  2.776633  0.2000
[4,]  -6.195870  4.835535  0.7155
[5,]  -1.264679  5.274504  0.0740

$crit
[1] 3.041755

>
VITA

Kristin E. Davidson lives in Oxford, Mississippi with her husband Gregg. She grew up in Tucson, Arizona, where she also attended the University of Arizona, earning a Bachelor of Science degree in Electrical Engineering. Upon graduation, she worked in the aerospace industry at Allied Signal as a Project Engineer, responsible for helping to design and manufacture fuel control systems for prop engines. After moving to Oxford, Mississippi and an eight year hiatus to raise her four children, she returned to graduate school, serving as an NSF Fellow, earning her Master of Science in Computer and Information Science. Following graduation, she has worked for the Computer Science department, beginning as a part-time Adjunct Instructor, and, since 2013, as a permanent full-time Instructor. She is a member of Phi Kappa Phi and Upsilon Pi Epsilon, honor society for computing and information disciplines.