RETENTION: COURSE COMPLETION RATES IN ONLINE DISTANCE LEARNING

Alana S. Phillips

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

December 2015

APPROVED:

Scott Warren, Major Professor
Michelle Wircenski, Committee Member
Demetria Ennis-Cole, Committee Member
Jeff M. Allen, Chair of the Department of Learning Technologies
Costas Tsatsoulis, Dean of the Toulouse Graduate School
Phillips, Alana S. *Retention: Course Completion Rates in Online Distance Learning.*
Doctor of Philosophy (Learning Technologies), December 2015, 118 pp., 12 tables, 5 figures, references, 162 titles.

Online courses in higher education have a reputation for having a lower course completion or retention rate than face-to-face courses. Much of this reputation is based upon anecdotal evidence, is outdated, or is on a small scale, such as a comparison of individual courses or programs of instruction. A causal-comparative analysis was conducted among 11 large, high research public universities. The universities were compared to each other to determine if differences existed between online and face-to-face course completion; undergraduate and graduation online course completion was analyzed for differences as well. The findings suggested the magnitude of the differences between online and face-to-face completions rates was small or negligible. The area which showed a higher magnitude of difference was in the comparison between undergraduate and graduate online course completion; the practical significance could be worth considering for educational purposes.
ACKNOWLEDGEMENTS

There are more people deserving of acknowledgment than I will be able to include here, so I will keep it brief. Thank you to my committee, of course. Scott Warren, my major professor, knew this was an area which needed research when I casually presented it along with a few other ideas. I am thankful for his suggestions and guidance during the journey and appreciate the way he allowed me to learn and form my own opinions. Demetrious Ennis-Cole and Mickey Wircenski, my committee members, were both supportive and encouraging throughout. Lynne Cox, Rebecca How, Pam Ponners, and Heather Robinson were a few of my friends who helped me by listening and giving excellent suggestions when I needed them.

My family did not complain too much when I quit cooking at the beginning of the program and helped out in so many other areas. Specifically, my husband Rick left me alone when I needed it, my daughter Taylor helped me with basic philosophical discussions, and my daughter Holly became my editor and mathematician. I am especially proud of the fact that my daughters and I all graduated from the University of North Texas together the same weekend in May 2015. As Lynne Cox and Scott Warren said in the systems class, everything is connected. Another connection came as I was scheduled to defend my dissertation in New Jersey. I missed the memorial service in Oklahoma for my grandma who went to be with Jesus at age 99. She was a teacher and would have understood. The memorial service was on the day my father, also a teacher, would have turned 79. I am quite sure they are a large part of the reason I decided to pursue a PhD.

Most importantly, I thank God for bringing me to a point in my life where I wanted to do the right thing instead of the thing I wanted to do. Fortunately, they can often be the same thing.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................................ iii

TABLE OF CONTENTS ............................................................................................................................ iv

LIST OF TABLES ................................................................................................................................ vii

LIST OF FIGURES ............................................................................................................................. viii

CHAPTER 1 INTRODUCTION ................................................................................................................. 1

Statement of the Problem ......................................................................................................................... 1

Purpose of the Study ................................................................................................................................. 2

Research Questions ................................................................................................................................. 2

Rationales .................................................................................................................................................. 3

Conceptual Framework ............................................................................................................................ 4

Research Method ..................................................................................................................................... 6

Operational Definitions ........................................................................................................................... 6

Limitations .............................................................................................................................................. 9

Delimitations .......................................................................................................................................... 9

Significance of the Study ......................................................................................................................... 9

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

CHAPTER 2 RELATED LITERATURE ................................................................................................. 11

History of Online Learning .................................................................................................................... 11

History of Learning Management Systems .......................................................................................... 12

Online Education Today ......................................................................................................................... 13

Retention in Education Settings .......................................................................................................... 14

Online Retention Rate Studies ............................................................................................................. 17
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparing Online to Face-To-Face Retention</td>
<td>18</td>
</tr>
<tr>
<td>Models of Online Retention or Persistence</td>
<td>30</td>
</tr>
<tr>
<td>Theoretical Framework</td>
<td>32</td>
</tr>
<tr>
<td>Traditional Student Retention Models</td>
<td>32</td>
</tr>
<tr>
<td>Non-Traditional Student Retention Model</td>
<td>33</td>
</tr>
<tr>
<td>Distance Education Retention Model</td>
<td>34</td>
</tr>
<tr>
<td>Online Student Retention Model</td>
<td>35</td>
</tr>
<tr>
<td>Summary</td>
<td>36</td>
</tr>
<tr>
<td>CHAPTER 3 METHODOLOGY</td>
<td>38</td>
</tr>
<tr>
<td>Paradigm or World View</td>
<td>38</td>
</tr>
<tr>
<td>Research Design</td>
<td>40</td>
</tr>
<tr>
<td>Research Questions</td>
<td>40</td>
</tr>
<tr>
<td>Population Sample</td>
<td>41</td>
</tr>
<tr>
<td>Data Collection</td>
<td>42</td>
</tr>
<tr>
<td>Variables</td>
<td>44</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>44</td>
</tr>
<tr>
<td>Reliability and Validity</td>
<td>57</td>
</tr>
<tr>
<td>Ethical Issues</td>
<td>57</td>
</tr>
<tr>
<td>Summary</td>
<td>57</td>
</tr>
<tr>
<td>CHAPTER 4 RESULTS</td>
<td>59</td>
</tr>
<tr>
<td>Research Questions</td>
<td>59</td>
</tr>
<tr>
<td>Descriptive Statistics</td>
<td>59</td>
</tr>
<tr>
<td>Inferential Statistics</td>
<td>64</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 3.1 Example Demonstrating Simpson’s Paradox</td>
<td>46</td>
</tr>
<tr>
<td>Table 4.1 Mean Overall Course Completion Rates for Fall 2013 and Spring 2014</td>
<td>60</td>
</tr>
<tr>
<td>Table 4.2 Descriptive Statistics for Overall Enrollments for Fall 2013 and Spring 2014</td>
<td>61</td>
</tr>
<tr>
<td>Table 4.3 University Enrollment and Completion Rates</td>
<td>62</td>
</tr>
<tr>
<td>Table 4.4 Online Completers and Non-Completers by University</td>
<td>65</td>
</tr>
<tr>
<td>Table 4.5 Online and Face-to-Face Course Completers by University</td>
<td>67</td>
</tr>
<tr>
<td>Table 4.6 Undergraduate and Graduate Online Course Completers by University</td>
<td>68</td>
</tr>
<tr>
<td>Table 4.7 Online and Face-to-Face Students: GLMM logistic parameter estimates</td>
<td>70</td>
</tr>
<tr>
<td>Table 4.8 Fixed experimenters with Random Intercept Model</td>
<td>71</td>
</tr>
<tr>
<td>Table 4.9 Online Students: GLMM logistic parameter estimates</td>
<td>72</td>
</tr>
<tr>
<td>Table 4.10 Fixed Effects with Random Intercept Model</td>
<td>73</td>
</tr>
<tr>
<td>Table 5.1 University Online Completion Rate in Descending Order</td>
<td>76</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Conceptual framework</td>
<td>5</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Composite persistence model</td>
<td>36</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Frequencies of online student enrollment</td>
<td>63</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>Frequencies of face-to-face student enrollment</td>
<td>63</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>Online and face-to-face completion rates</td>
<td>64</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION
Statement of the Problem

Retention for online courses in higher education has a reputation for being lower than in face-to-face courses. Previous retention research has been small in scale, focused mainly on individual courses or community college settings; some of it is outdated. While Allen and Seaman (2015) reported that chief academic officers perceive online retention as low, a review of literature presents a research deficiency regarding support for an online retention problem. Additionally, many of the previous studies were conducted before distance learning evolved to include online instruction and current Internet technologies became the primary medium for conducting distance learning. Most retention studies now examine the effectiveness of models used to improve online retention or identify other reasons online students drop out of classes.

The National Center for Education Statistics which tracks retention does not currently distinguish between online and face-to-face retention (“Digest of Education Statistics,” 2014). Retention rates, or percentages, come from reports using anecdotal examples or individual university rates, with no national rates available (Berge & Huang, 2004; Choi, Lee, Jung, & Latchem, 2013; Lawlor, 2007). As Clark (1983) challenged the research in media comparison studies, there should be an online retention discussion to challenge the research, noting that “consistent evidence is found for the generalization that there are no learning benefits to be gained from employing any specific medium to deliver instruction” (Clark, 1983, p. 445) and that this should therefore also apply to online instruction.

As stated, much of the existing research was conducted on the reasons students drop out of online courses or programs (Moody, 2004; Nash, 2005; Tyler-Smith, 2006) and models or
ways to improve retention, persistence or attrition (Angelino & Natvig, 2009; Berge & Huang, 2004; Simpson, 2003). The reasons given for the lower retention rate include different demographics between online and traditional or face-to-face students, such as age or previous college experience (Diaz, 2002; Patterson & McFadden, 2009) and issues with course development or student preparation (Moody, 2004).

Purpose of the Study

This study compared retention, also known as course completion, between online delivered and face-to-face courses at select four-year public universities. The goal of this study was not to prove retention is or is not lower in online learning than in face-to-face courses; rather it was to improve and expand the existing research which describes a difference in retention between online learning and face-to-face learning among institutions.

Research Questions

This study was driven by a need to either substantiate or dispel the idea that retention is lower in online courses. Further, other research questions are appropriate as well, including comparing institutions and student classification across the Integrated Postsecondary Education Data System (IPEDS) geographical regions ("College & Careers," n.d.).

The research questions for this study were:

1. What is the difference among select four-year universities’ course completion rates of online distance learning courses across the regions of the United States?
2. What is the difference in online distance learning course completion and face-to-face course completion among select four-year universities across the regions of the United States?
3. What is the difference in online distance learning course completion between undergraduate and graduate students at select four-year universities across the regions of the United States?
These questions were predicated on the assumption that institutions or agencies would be able to provide data requested using the criteria requested, and that universities review their own completion rates with a view towards deciding if the data made sense.

Rationales

It was determined from the literature that additional research was needed to substantiate or debunk the claim (Carr, 2000; Frankola, 2001) that online retention is lower than face-to-face retention. Continuing to perpetuate the idea that online learning has a higher dropout rate can be a self-fulfilling prophecy; online students may be predisposed to dropping out if they think everyone else drops out of online courses more often than face-to-face courses. Merton (1948) coined the term self-fulfilling prophecy based on the Thomas theorem: “If men define situations as real, they are real in their consequences” (p. 193).

Rosenthal and Jacobson (1968) took this idea into classrooms to see if teachers could influence students based on their perceptions of student ability, finding statistically significant results with the younger students, which were first and second graders. This study generated controversy; however, Rosenthal’s (1976) meta-analysis found 35% of 311 studies showed statistically significant results as well. Rosenthal termed this the experimenter expectancy effect, in which a subject in an experiment might be more likely to deliver what the experimenter expects to be the response or result. The Hawthorne effect is similar, where the subjects’ behaviors can change because they know they are research subjects, they know the hypothesis, or because they are getting extra attention (Gall, Gall, & Borg, 2007; Gillespie, 1991). Another related consideration is the novelty effect, which can occur when subjects perceive and then act differently than they would normally (Gravetter, F., & Forzano, L. A., 2011) and can be related to the introduction of new technology (Clark, 1983).
If there is a “great deal of evidence that teachers’ expectations can function as self-filling prophecies” (Brophy, 1983, p. 632), this study and others like it should be pursued. Jussim and Harber (2005) revisited this controversial area of study over self-fulfilling prophecies, and their recommendation could also apply to online retention beliefs: We should be “relying on the actual results of the relevant studies, rather than on a consensus of scholarly opinion” (p. 132).

**Conceptual Framework**

A conceptual framework is used in different ways which leads to at least three different definitions (Ravitch & Riggan, 2012). The definition used here is that a conceptual framework is “an argument about why the topic one wishes to study matters, and why the means proposed to study it are appropriate and rigorous” (Ravitch & Riggan, 2012, p. 7).

In this sense, the purpose of the study was to compare retention or course completion in online learning and face-to-face learning at select four-year public universities. The larger purpose of the study was to add to existing research on the difference between online and face-to-face course completion or retention rates and to function as exploratory research.

Stebbins’ (2001) discussion of exploratory research helps describe the current state of the comparison between online and face-to-face course completion rates:

In general, exploration is the preferred methodological approach under at least three conditions: when a group, process, activity, or situation has received little or no systematic empirical scrutiny, has been largely examined using prediction and control rather than flexibility and open-mindedness, or has grown to maturity . . . but has changed so much along the way that it begs to be explored anew. (p. 9)

This statement situates this study in the last two words: explored anew. Technology has changed over the past ten year and previous statistical analyses have had shortcomings (Kline, 2013), as will be discussed in further detail in the literature review.
This study’s methods are appropriate because the two most cited articles (Carr, 2001; Frankola, 2001) and others have not shown appropriate or rigorous methods of analysis. The first two articles consisted of anecdotal evidence, while the other references often failed to report effect size. Those shortcomings found in the literature helped guide the methods in this study.

The conceptual framework is shown graphically in Figure 1. Much of the basis for

![Conceptual Framework Diagram]

**Figure 1.1.** The conceptual framework shows that changes in people and technology have driven online instruction to evolve from distance learning and mature into online instruction today. Empirical analysis of online course completion should be at the center of the growth.

claiming there is an online retention problem was built on anecdotal data collected outside of the research community. Empirical research should be part of the evolution and growth and is an important step in online instruction.
Research Method

The research employed causal-comparative analysis, or criterion-group design, which is non-experimental analysis on existing groups, in this case online and face-to-face students (Fraenkel, Wallen, & Hyun, 2011). Causal-comparative analysis can be an alternative to an experiment, and then the results could be used as the basis for further research depending upon the original causal-comparative analysis (Fraenkel et al., 2011). The quantitative analysis reported here is the first phase of a mixed methods study. Future, follow-on research is planned using extreme case sampling of the universities with the highest and lowest online course completion rates (Fraenkel et al., 2011; Harrits, 2011) and accompanying qualitative analysis (Ames, Duke, Moore, & Cunradi, 2009; Creswell & Clark, 2010).

Operational Definitions

Attrition consists of the failures, withdrawals, or dropouts subtracted from the number of students who started the course as of the census date. It will refer only to the course, not to the program of study or the institution.

Blended course is a course with 30% to 79% of the instruction online and the rest in face-to-face meetings; it is the same as a hybrid course (Allen & Seaman, 2015; Robinson, Phillips, Sheffield, & Moore, 2015). Moore and Kearsley (2011) add the distinction that blended learning is “complemented by technology”; distance education or online learning requires technology (p. 3).

Census date is the date at the beginning of the course period when the institution reports enrollment data to the governing body.
Course completion can be defined differently by institutions, but generally is receiving a passing grade in a course. Students who completed a course unsuccessfully are counted in the failure category, for example, those who earned a grade of D or F.

Distance education is instruction where the student and instructor are not together at either the same time or the same place, or both, which requires technology for communication (Moore & Kearsley, 2011).

Drop out can be measured as a student who fails to complete a course, for the purpose of this study. This is normally referred to as W/D or W/F. A student makes the decision to drop out but it is considered the same as a failure or withdrawal for the purposes of this study. This does not refer to a student who completes a course but departs the institution afterwards, whether completing their goal or not.

Failure can be measured as a student who receives a failing grade (F) or other measure according to the institution and has not successfully completed the course. An incomplete (I) can become a failure if the work is not completed in 12 months. Some institutions or programs may count a D as a failure (Diaz, 2000).

Face-to-face course is a course where there is no technological content or up to 29% of the content may be delivered online. These can be considered either traditional or web-facilitated courses (Allen & Seaman, 2015).

A hybrid course is a course with 30% to 79% of the instruction online and the rest in face-to-face meetings; it is the same as a blended course (Allen & Seaman, 2015; Robinson et al., 2015).

Incomplete can be defined by the institution, but is normally a student who has not completed the course and has 12 months to do so.
Online course is a course where the instructor normally delivers over 80% (Allen & Seaman, 2015) of the “content, instruction, and materials over the Internet and the student attends class within this online classroom” (Robinson et al., 2015, p. 58). Moore and Kearsley (2011) emphasized that an online course meets at a different time or location or both, so the technology is not what makes it an online course, it is how the instructor meets with the students.

Program refers to the courses, hours, and other requirements that a student progresses through until completion. Not all students are in a program of study.

Proportion is “an expression in which the numerator is always included in the denominator, and the base is equal to 100. Therefore, a proportion is always expressed as a percent” (Peavy, Dyal, & Eddins, 1983, p. 2).

Rate “measures the probability of occurrence of some particular event” (Peavy et al., 1983, p. 1).

Ratio is an “expression of the relationship between a numerator and denominator which may involve either an interval in time or may be instantaneous in time” (Peavy et al., 1983, p. 2). The numerator is not necessarily a portion of the denominator.

Retention is indicated by a student successfully completing a course and being able to proceed on to the next goal, for the purpose of this study. The goal could be another course or set of courses, graduation, or receipt of a certificate. Using Mortenson’s (2005) definitions, a series of course completion rates, graduation rates, dropout rates, and transfer rates together become the retention rate.

Withdrawal is the same as drop out for the purpose of this study. Withdrawal refers to a student who fails to complete a course and does not refer to whether a student remains at the
university or in a program of study. The instructor or the student may make the withdrawal decision.

Limitations

It could be argued that it is inappropriate to compare face-to-face course completion or retention with online course completion (Fraenkel, Wallen, & Hyun, 2011). The courses commonly attract different student audiences (Distance Education and Training Council, 2007; Moore & Kearsley, 2011).

Secondary data was drawn from different sources and there is a chance the data providers interpreted the request for data differently than intended and a university was represented incorrectly.

Delimitations

Delimitations included:

1. The data was only drawn from select large four-year universities. This design could be considered an asset to the study though, as it will make the comparison groups more homogenous (Fraenkel et al., 2011).

2. Course completion information was used instead of retention as defined by IPEDS.

Significance of the Study

The study adds to the existing research on retention in online learning given the absence of a national database. Studies to improve retention in online learning and face-to-face learning are important, but adding to the existing isolated and anecdotal information about online retention is key to furthering research. Given the conclusion Clark (1983) reached – that the media does not matter – it is likely the same conclusion would be reached in a broad study of student completion rates in online learning. Other factors may be more important than the media.
“Education, not retention, should be the goal of institutional retention programs” (Tinto, 1987, p. 140).

Summary

This chapter provided a summary of the purpose of study, which was to compare retention or course completion in online courses with face-to-face courses. The chapter included the rationale and significance. Briefly, this is to improve upon existing statements used in online education research about the quality of online education when compared to face-to-face education. Course completion or retention is often used as a measure of quality but is not based upon a large body of empirical research. This can lead to misperceptions of online instruction, which can in turn lead to lower completion rates. The literature review follows in the next chapter and is intended to explain in more detail how the misperceptions occur in research and discuss the existing research which compares online and face-to-face course completion.
CHAPTER 2

RELATED LITERATURE

The literature starts with a history of online learning, covers online retention rate studies within the last ten years, then includes a discussion about models of online student retention. Online learning has evolved over the past ten years along with advances in technology. Negative perceptions about educational quality that have been held since distance learning began still persist today (Larreamendy-Joerns & Leinhardt, 2006).

History of Online Learning

Current perceptions of online learning may have historical roots. The history of online learning is reviewed first, followed by a history of learning management systems. Both may influence current perceptions of online learning.

Online learning evolved from distance learning. Distance education started with correspondence courses in the 1840s in Great Britain and the 1880s in the United States, where content was delivered by mail. Moore and Kearsley (2011) called this first generation distance education. Thirty-six years after the first commercial correspondence school opened in 1891, correspondence schools enrolled about two million students, four times the number of students in traditional colleges and professional schools (Fisher, 1927).

The second generation included courses taught by radio and television broadcast starting in the 1920s. Distance education’s third generation began in the 1960s with integrated media programs that combined print media, broadcasts, audio recordings, telephone conferences, home experiment kits, and local library resources. One version was called the Articulated Instructional Media Project (AIM) and led to Great Britain’s development of The Open University (OU). The
OU started emphasizing learner support by offering tutoring and counseling (Moore & Kearsley, 2011).

The fourth generation was marked by teleconferencing within groups. This started with audio-conferencing in the 1970s, which enabled distance learning students to communicate synchronously like those in traditional classrooms. Satellite technology allowed distance education providers to transmit interactive video-conferences in the 1980s (Moore & Kearsley, 2011).

Moore and Kearsley (2011) labeled computer and Internet-based classes as the fifth generation in distance education. Universities started offering web-based courses and programs in the 1990s. Courses which were originally based on the independent study method of correspondence courses have been using Web 2.0 technologies such as social networking to become virtual classrooms and instead meet synchronously (Moore & Kearsley, 2011). The Web 2.0 technologies started in the early 2000s and can be described as second generation Internet tools which allow users to create their own content and collaborate easily (den Exter, Rowe, Boyd, & Lloyd, 2012). Web 2.0 tools such as discussion forums, wikis, and blogs may be implemented through a learning management system, or LMS (den Exter et al., 2012).

*History of Learning Management Systems.*

Baker (1971) traced the predecessors of today’s learning management systems (LMS) back to 1959, as computer-assisted instruction systems. Other terms for LMS can be learner management system, course or content management system, or integrated learning system (Moore & Kearsley, 2011). Baker (1971) noted computer-based instructional management systems then had four major functions: test scoring, diagnosing, prescribing, and reporting.
The first LMSs started with professors using small tools in their own classes and have evolved into large enterprise systems (Morgan, 2003). Early LMSs were accessed through a mainframe terminal and later the personal computer using a local area network. A university in New Zealand believes they released the first web-based LMS in early 1996 (Sheridan, Gardner, & White, 2002).

Most LMSs in the early 2000s were course-centered rather than student-centered, which meant they lacked personalization, collaboration, and integration for the student (Graebner, 2000). Changes in the LMS can be tied to Web 2.0 technologies. Vossen and Hagemann (2007) looked at the Web 2.0 developments in terms of streams: application, technology, and user participation and contribution, or socialization.

Broadband usage started overtaking dialup usage in 2004 (Organization for Economic Co-Operation and Development, 2011), which enabled more people to use new applications and services and helped bring continued improvements in multimedia technology (Rath, 2000). Powell and Keen (2006) expected that “once we have enough bandwidth, after all, students will be able to gather together in virtual classes all at once, wherever they are, by sitting in front of their home computers” (p. 294).

**Online Education Today**

Online education has roots in both distance education and computer-assisted instruction. Some in academia view online education as a threat to education while others view it as an opportunity to “overcome the limitations of traditional classroom instruction” (Larreamendy-Joerns & Leinhardt, 2006, p. 572). Instructional quality is an acceptance issue that distance education and computer-assisted instruction have had that can be expected to continue for online
education. Retention, attrition, or success rate is viewed as one indicator of instructional quality (Larreamendy-Joerns & Leinhardt, 2006).

Retention in Education Settings

Using a definition from Merriam-Webster, retention is “the act of keeping someone or something” ("Retention," 2014, para. 1). Another definition is that retention is “the ability of a particular college or university to successfully graduate the students that initially enroll at that institution” (Berger & Lyon, 2005, p. 3). Mortenson (2005) explained that persistence or retention is measured through a “series of status-to-status ratios… called transition ratios, persistence rates, retention rates, completion rates, cohort survival rates, or graduation rates… and other measures of change in student status” (p. 31).

The National Center for Education Statistics (NCES) is the U.S. agency responsible for collecting and analyzing education data ("About us," n.d.). The NCES database for post-secondary education is IPEDS ("Integrated Postsecondary," n.d.). Institutional level data is collected from colleges, universities, and technical and vocational institutions into IPEDS in areas which include enrollment, program completion, and graduation rates. The enrollment data has been used to calculate retention rates since 2003 (Cunningham, Milam, & Statham, 2005).

The IPEDS definition of retention refers to freshmen and does not distinguish between face-to-face and online students. For the purpose of studying either delivery method separately, the following retention definition will not suffice:

A measure of the rate at which students persist in their educational program at an institution, expressed as a percentage. For four-year institutions, this is the percentage of first-time bachelors (or equivalent) degree-seeking undergraduates from the previous fall who are again enrolled in the current fall. For all other institutions this is the percentage of first-time degree/certificate-seeking students from the previous fall who either re-enrolled or successfully completed their program by the current fall ("Glossary," n.d., letter R).
The University of California, Merced, added to the IPEDS definition of retention rate to include “a measure of academic progression of a group of students from one period of time to the next” ("Terms," 2009). Simonson’s (2010) definition for online instruction is more specific: “the likelihood that online students … will persevere and continue the course, certification, or degree program to completion” (p. 91).

Similar terms are completion rate, course completion ratio, successful course completion ratio, or pace of completion which is used by community colleges (Hagedorn, 2005) and to verify and maintain financial aid eligibility (U.S. Department of Education, 2014). Normally considered from a student’s perspective, these terms refer to the percentage of classes completed successfully. This definition sounds like persistence (discussed in the next section), but the Research & Planning Group for California Community Colleges ("Institutional Research," 2011) changed their definition of persistence to become retention and their retention definition to become course completion rate.

The Center for the Study of College Student Retention suggested the online retention definition be modified from the standard retention definition to be: “a student is retained in a distributed learning course and/or program if he/she is making satisfactory progress towards a personal and/or educational objective consistent with the college’s mission” ("Retention definitions," n.d., para. 7). Additionally, the center recommended the course retention definition be “the number of students enrolled in each credit course after the course census date and the number of students who successfully complete the course with an A-D grade at the end of the term” ("Retention definitions," n.d., para. 5).

A study of online course retention at Christopher Newport University (Ridley, Miller, & Williams, 1995) used the term course retention and defined it as the “comparison between online
and classroom courses in students’ persistence, after enrollment, through the course to completion” (p. 2) and online retention, which is “the tendency of students to enroll in other online courses after having enrolled in one course” (p. 2).

To calculate retention rate, institutions may use a baseline number starting with students who register, start a class, or start submitting assignments (Kember, 1995; Simpson, 2003). The retention starting baseline can also be date referenced. For example, IPEDS defines the official fall reporting date as “the date (in the fall) on which an institution must report fall enrollment data to either the state, its board of trustees or governing board, or some other external governing body” (Broyles, 1995, p. 71).

For the purpose of this study, course completion rate was used instead of retention in order to refer to the students who completed an online class. Limiting retention rates to entire cohorts or programs would have failed to capture online courses which were not taken as part of an entire program. The baseline was students who were enrolled as of the date at the beginning of the course period when the institution reports enrollment data to the governing body, frequently called the census date. Students were counted as enrolling in the course if they received a grade of any kind which reflected a completion status, or incompletion status for a withdrawal, on their transcript. An incomplete not resolved within 12 months counted as a withdrawal.

*Persistence, Attrition, and Dropout*

Hagedorn (2005) explained that even though the terms retention and persistence are used interchangeably, “institutions retain and students persist” (p. 92); retention should be used as an institutional measurement and persistence as a student measurement. Hagedorn (2005) also stated that “attrition is the diminution in numbers of students resulting from lower student
retention” (p. 92) and a dropout is “anyone who leaves college prior to graduation”. She also explained that a dropout from one school can be a transfer to another school or can return to the original school later and is no longer a dropout.

Tinto (1975) cautioned against grouping all dropouts together instead of distinguishing between academic dropouts and voluntary withdrawals, or distinguishing between permanent and temporary dropouts or transfers to other institutions. “If the leaver does not define his/her own behavior as representing a form of failure, neither should the institution” (Tinto, 1987, p. 132). Tinto (1987) considered students a dropout once they consider their leaving as a failure to reach their intended goal – then it also became a failure of the institution to help the students achieve their original goals. Phipps, Merisotis, Harvey, and O’Brien (2000) expressed the perspective that “distance learning studies should clearly indicate the beginning and ending number of students in a course to ensure that dropouts are not excluded from the analysis.”

Attrition, according to Martinez (2003) is a “decrease in the number of learners or students engaged in some course of study. The course of study might be a degree plan, or it might simply be a standalone online course” (p. 2). In the past, a certain level of attrition was considered a positive indicator for more selective institutions. Most colleges focused on getting new students rather than keeping current students (Berger & Lyon, 2005). Hossler and Bean (1990) considered the terms attrition and retention as interchangeable, even though they mean the opposite. The authors used models as an example, stating that attrition models are also retention models.

Online Retention Rate Studies

In the literature, the references to online learning having lower retention rates than face-to-face learning came from the same few studies or articles. Diaz (2002) explained that “the
notion that more students will drop out of online classes than traditional face-to-face classes enjoys the widespread acceptance usually reserved for scientific precepts” (para. 1). Lawlor (2007) referred to this claim of lower online retention as a perception. Chyung (2001) and Kember (1995) suggested universities may not want to publicize their online retention rates. Also, Kember (1995) cautioned that retention data that is available might not be measured the same among institutions and is not appropriate for comparison.

Comparing Online to Face-To-Face Retention

Most studies of retention rate or course completion were conducted at individual community colleges or among individual or multiple course offerings. Many of the studies published in the past ten years were from dissertations and few reported effect size. The following review summarizes findings of the studies. Each summary includes the results of the study followed by the perception of online retention which the author(s) presented in the literature review.

Atchley, Wingenbach, and Akers’s (2013) study of one university in Texas compared course completion among online and face-to-face students over a five-year period. The study was a causal-comparative analysis of retention determined by course completion and performance between online and traditional courses from Fall 2004 through Spring 2009. The study included all enrolled students; the article does not specify if students were undergraduate or graduate level. Atchley et al. (2013) found a statistically significant difference between course completion rate, with 93.3% online and 95.6% for face-to-face courses (n = 5,778). Effect size was not reported. The literature review referenced Carr (2000), Roach (2002) and Waschull (2001) and stated that research comparing online and traditional course completion rates led to mixed results.
Norman (2013) compared retention rates, course grades, and success rates between online and face-to-face students in Tennessee who had recently completed remedial courses. The dissertation was a mixed methods causal-comparative study between a community college’s eight English composition courses and a medium sized four-year public university’s algebra class. The quantitative analysis came from student grades and demographic surveys; qualitative analysis was from emailed interview responses. There was a statistically significant difference between sample populations of students with at least one prior remedial course, with student retention in the online/hybrid classes at 56.8% (n = 46) and face-to-face at 43.2% (n = 35). Effect size was not reported. The introduction referenced Carr (2000) to indicate online retention rates were lower than in traditional courses.

Tornsaufer (2013) used retention data from California’s Data Mart, analyzing results submitted by 112 community colleges. The dissertation was a quantitative analysis comparing ethnicity and gender among the course delivery methods of online asynchronous, online synchronous, and traditional face-to-face courses. The statewide results showed Fall 2011 retention rates differed, with asynchronous at 77.7% (n = 363,338), synchronous at 80.7% (n = 34,430) and face-to-face at 84.9% (n = 4,074,018). Tornsaufer (2013) did not report an analysis between the course delivery methods—only between ethnic groups and genders—but did detect a statistically significant difference between male student retention and asynchronous and synchronous instruction. There was no significant difference between female student retention and the online delivery method. There was a statistically significant difference in retention between both genders for online course delivery and face-to-face delivery; effect size, calculated as $\eta^2$, was large (.245). The introduction stated there was a need to increase online retention, but the reference used to substantiate an online retention problem did not mention online retention.
Xu and Jaggars (2011) conducted a quantitative study comparing entry level online course enrollment among Virginia’s 23 community colleges from 2004 through 2008. The researchers used propensity matching to compare course persistence and performance between online and face-to-face students who took entry level English or math courses, both across and within schools. English course online attrition was 19% and face-to-face 10%. Math course online attrition was 25% and 12% face-to-face. Xu and Jaggars (2011) calculated the odds ratio as effect size and determined online English students were 2.37 times more likely to dropout than face-to-face students using the full sample, 1.93 times more likely using within schools, and 2.27 times more likely to dropout using across-school propensity matching. Math students in online courses were 2.93, 2.70, and 2.92 times more likely to dropout, respectively. The last year of the study resulted in online and face-to-face attrition rates in English which were similar, but once the researchers adjusted using the propensity matching model, results for that last year did not change significantly from the earlier years of the analysis. The authors also suggested that improving technologies from 2004 to 2008 did not impact online success (Xu & Jaggars, 2011). Xu and Jaggars (2011) stated that most research comparing online and traditional education was in areas other than course completion, such as differences in learning outcomes, and did not make a claim as to what the literature said about online retention.

Long-Goding (2011) conducted a quantitative analysis of online retention before and after a treatment was applied. The treatment consisted of professional development training for advisors. Part of the analysis in this dissertation consisted of comparing face-to-face and online retention, or course completion, for the Spring 2010 semester at a community college. Retention rates between online and face-to-face differed by 8% and a Wilcoxon Signed Rank test showed the difference was statistically significant. Long-Goding (2011) observed that the retention gap
had improved from 24% up to 8% over the past nine years that the Massachusetts college had offered online courses. Effect size was not reported. In the first chapter, the authors acknowledge the debate in online retention and listed Roach (2002) and several other secondary references to substantiate the claim that retention was “frequently lower” in online education (Long-Goding, 2011, p. 4).

Patterson & McFadden (2009) compared attrition in two programs of online and face-to-face master’s students (n = 640) at a large research university in the Southeast. In the quantitative study, the researchers analyzed enrollment records from Fall 2002 through Fall 2004 and conducted logistic regression analysis which indicated course format and student age had a “significant unique effect” (p. 10) on combined dropouts. Demographic and academic variables considered were found not to have a statistically significant effect on online student dropout rate. Patterson & McFadden (2009) found a statistically significant difference in the dropout rate between online and face-to-face students in both programs. In the time period studied, 11% of traditional students dropped out while 43% of online students dropped out of the MBA program (n = 516). Effect size was reported as an odds ratio of 5.9. The second program, communication sciences and disorders, had 4% of face-to-face students and 23.5% of online students drop out (n = 112). The odds ratio was stated as seven in the discussion but not specifically reported. Patterson and McFadden (2009) claimed that attrition is “often much higher” (p. 2) in online courses and used secondary references such as Rovai (2003) and other articles to substantiate this.

Frydenberg (2007) compared enrollment and dropouts from Spring 2004 through Winter 2006 of adults in continuing professional education courses and did not find a statistically significant difference between online and face-to-face students. This quantitative study
conducted at the University of California Irvine determined online students (n = 183) dropped at 8%, the same rate as face-to-face students (n = 2,512).

Caldwell’s (2006) study of undergraduate students in an introductory computer science programming course compared online, face-to-face, and web-enhanced course completion, performance, motivation, and satisfaction at Winston-Salem State University. Students were randomly assigned into the course format for the study. Five students enrolled but never attended classes (both face-to-face and online). The remaining students (n = 55) all completed the classes.

In the literature review, Caldwell (2006) indicated research had not explained attrition, included Phipps and Merisotis (1999) as a secondary reference and included their references, along with two other studies with mixed results. Then she started the next paragraph with “the high dropout” as if it were a fact (p. 37).

Nelson (2006) compared online student retention with comparable courses consisting of face-to-face students. The quantitative study at one campus of the Delaware Community and Technical College was conducted using the causal-comparative method to consider enrollments from Fall 2003 through Fall 2005 (excluding summer semesters). Nelson’s (2006) dissertation presented results of a chi-square analysis which indicated there was a statistically significant difference between the retention rates of online (n = 1,877) and face-to-face students (n = 3,308); online retention was 77% and face-to-face was 81%. Effect size was not reported. Further analysis by course subject indicated a statistically significant difference in retention between course formats for criminal justice and psychology students only. There was also a statistically significant difference between the number of males dropping out of online courses and face-to-face courses.
Paden (2006) compared retention in online, blended, and face-to-face course sections of an undergraduate introductory math course at a large private university. Retention was defined as completing the course. The quantitative study, Paden’s (2006) dissertation, used the causal-comparative method to analyze grades and retention records from January 2002 until December 2004. Retention for the face-to-face class was 91.9% (n = 40,179), online was 85.6% (n = 38,056), and blended was 88.2% (n = 1,310). A chi-square test indicated the differences were statistically significant. Effect size was not reported. In his introduction, Paden (2006) discussed the disconnect between beliefs and research in comparing online and face-to-face retention.

Del Negro (2005) studied 24 sections of 7 courses to compare online and face-to-face student performance. The dissertation looked at several performance-related factors in the comparison. Results showed that male online students dropped out at a higher rate than traditional face-to-face students with statistical significance. Additionally, chi-square analysis showed no significant differences existed between overall online and face-to-face course failure rates in the study while a significant difference was found in the withdrawal rates. Effect size was not reported. Del Negro (2005) did not make a claim about online retention or completion rates in the literature review.

McLaren (2004) compared online (n = 152) and face-to-face (n = 139) students in a university undergraduate business statistics class that she taught from Fall 2000 through Spring 2003. A chi-square test for independence ($\chi^2=51.701$, df = 2, p < .005) indicated a statistically significant difference in student persistence between the two classroom modes. Effect size was not reported. Seventy-one students dropped or failed to complete the online classes; 12 students dropped or failed to complete the face-to-face classes. McLaren (2004) discussed the debate on
comparing online and face-to-face courses and then referenced both Carr (2000) and Frankola (2001) when discussing persistence specifically.

Of the twelve preceding studies, eight showed a statistically significant difference between online and face-to-face retention, two showed no difference, and two showed a difference between male and female retention; three of these reported effect size. A common thread appearing in the literature reviews was the perception that online learning retention is lower than face-to-face. This common thread is explored in the section about common references which appears right after the meta-analysis.

A Meta-Analysis

A meta-analysis conducted before the ten-year time period is worth considering in the review. Bernard et al. (2004) conducted a meta-analysis of literature comparing distance and classroom education from 1985 to 2002. Retention, achievement, and attitude were analyzed in a quantitative synthesis of 232 studies. Distance education courses analyzed were considered third, fourth, and fifth generation (Moore & Kearsley, 2011), as discussed in the history of online learning at the beginning of this chapter. Retention in this study was measured by completing a course. Effect sizes for the studies in Bernard et al.’s (2004) meta-analysis were aggregated and the population for the retention analysis was truncated from n = 57,916,029 to n = 3,735,050 to compensate for the extreme effect. A distribution of effect sizes showed the primary mode was at zero; the secondary mode was at -.38. Overall, retention in classroom instruction had a “very small but significant effect” (p. 405) over distance education (g+ = -.0573).

Bernard’s review of literature discussed previous meta-analyses, most of which found no significant results. Bernard et al. (2004) also referred to Clark’s (1983, 2000) position on discouraging comparative studies based on media; Bernard et al. (2004) explained that
comparing distance to face-to-face education is not the same as a study comparing uses of different media. Distance education includes other factors such as teacher proximity to the learner and methods of communication.

**Common References for Low Online Retention Rates**

Most studies about retention have focused on explaining causes for retention or on improving retention in online learning. A large number of studies published in journals have used the same one, two, or three references to substantiate low online retention rates. The following are the most commonly referenced publications. Each reference is discussed along with the number of times the article was referenced according to Google Scholar. These references are not included in the literature review because they were over 10 years old.

Carr’s (2000) newspaper article provided anecdotal information from several individuals interviewed and restated course completion rates provided by two community colleges and two universities in three states. An administrator at a Dallas County Community College educational technology center was quoted as saying course completion rates for online courses were 11% to 15% lower than for face-to-face courses. Tyler Junior College’s course completion rates for that fall were reported as 58% for Internet courses and 71% for face-to-face courses. The University of Central Florida was reported to have a 9% withdrawal rate for 47 web-based courses and a 5% withdrawal rate for face-to-face courses in the same subjects during the fall of 1998. The University of California at Los Angeles was listed as having 50-60% completion rates the first few quarters after they started offering online courses; the rate was up to 87% for the previous eight quarters. Carr (2000) interviewed a student at Tyler Junior College who dropped an online course; the student said the instructor was not prepared. Carr (2000) interviewed a few other students who shared their personal reasons for dropping an online course, as well as listing
generalities without including sources, such as saying course completion rates for online courses were usually 10-20 percent lower than for face-to-face courses. Google Scholar reported 866 references to this article.

Frankola’s (2001) magazine article is similar to the Carr (2000) magazine article in that it used anecdotal evidence. It was cited in 339 articles according to Google Scholar.

Other studies exist which report retention or persistence in individual institutions or classes, but they are not referenced as often as the preceding two newspaper articles. The third most cited article is Waschull (2001). She conducted research on her own course sections which she taught as both online and face-to-face. Students self-selected into the class format or could enroll in different sections taught by other instructors. In the introductory psychology course, 33 students enrolled in either an online (n = 14) or face-to-face (n = 19) format. One online (7%) and three face-to-face (15%) students dropped the course, which was not statistically significant (Waschull, 2001). Google Scholar listed this article as being cited 136 times.

Diaz’s (2000) dissertation compared community college students (n = 96) in an online health education class to students (n = 135) in the face-to-face class. Academic success between online and face-to-face students was one of the comparisons. Academic success was defined as completing the course with a C or better. Distance students received twice as many A grades (25%, 12.9%) but the attrition was nearly double (13.5%, 7.2%). Diaz (2000) concluded that the online students were “somewhat more successful” than face-to-face students in the study (p. 78). Diaz (2000) was referenced 78 times according to Google Scholar.

Brady (2001) personally tracked her own university English courses and found a higher percentage of dropouts in the online versions. One class size was 22 students; the author does not state the other class sizes. The author followed two semesters of face-to-face classroom (4%
withdrawal rate) and four semesters of online (14% withdrawal rate) offerings for one course and four semesters of face-to-face classroom (5% withdrawal rate) and eight semesters of online (13% withdrawal rate) offerings for the other course. Based on her experiences, Brady (2001) stated she had “retention problems as reflected in withdrawal rates for the distance version of the course that are almost three times higher than they are for the traditional classroom version” p. 348). Brady (2001) was cited 52 times according to Google Scholar.

Roach’s (2002) newsmagazine article included information from two private colleges and one community college substantiating their experience with higher retention rates among online students when compared to face-to-face students. Roach (2002) explained that “research on retention is scarce mainly because of the newness of online education, but individual schools and organizations are reporting that their online programs have as high or higher rates of retention as their traditional classroom offerings” (p. 23). Google Scholar listed Roach (2002) as being cited 33 times.

In another study, Ridley et al. (1995) tracked the first two semesters—eight fall courses and twelve spring courses—of a new online delivery program, and students (n = 507) dropped courses at a rate three times higher than the overall university rate; the dropout rate declined from 30% to 25% in the second semester. Google Scholar listed this study as being referenced by seven other articles.

Johnson’s (2003) study compared community college students in online courses to those in face-to-face courses and found face-to-face students had a 19% higher course completion rate than online students. This dissertation did not specify how students were selected for the study and counted students who dropped a course in the first ten days as not completing a course. Online students (N = 305) were enrolled in 43 different academic and technical courses, while
face-to-face students (n = 149) were enrolled in 15 different courses for the spring of 2002. Thirty-eight percent of online students did not complete a course, and 16.8% of face-to-face students did not complete a course. A t-test was used to compare means but was not included in the dissertation (Johnson, 2003). There were seven references that cited this study according to Google Scholar.

Of the eight references most used in the literature review of online retention rate studies, the top two presented anecdotal information to substantiate online retention as being higher than face-to-face retention. The third most referenced study found no statistical difference between the delivery methods. The remaining five which were referenced less often found a difference, with online delivery retention being lower than face-to-face.

Other Retention Research in Online Education

Some researchers have questioned the existence of data or of an online retention issue (Boston & Ice, 2011; Diaz, 2002; Tyler-Smith, 2006), yet a cursory review of literature shows research to solve a retention problem has been more prolific than research to confirm a problem. Other researchers have disagreed over the quality of research in distance education, such as whether studies should use a control group and randomize subjects (Brown & Wack, 1999; Phipps & Merisotis, 1999), which could place a higher burden of proof on distance education than on other educational research (Brown & Wack, 1999).

Lockee, Burton, and Cross (1999) observed that distance education appeared to be growing faster than research studies. McIsaac and Gunawardena (1996) explained that “research has reflected rather than driven practice” (p. 403). They stated that researchers have studied areas which distance education administrators are interested in, including attrition as well as instructional material design, technology use, and cost effectiveness (McIsaac & Gunawardena,
Explaining Causes of Attrition or Dropout in Online Learning

Phillips (2013) studied non-traditional students at a four-year Midwestern university to compare retention differences between online, hybrid, and traditional students. This dissertation was a quantitative study using survey instruments to measure student retention factors. Phillips (2013) concluded there was a statistically significant difference in retention factors between online, hybrid, and face-to-face students, and that online delivery “was more effective at promoting factors responsible for the retention” of nontraditional students than hybrid and face-to-face delivery. Phillips (2013) referenced Johnson (2003) to substantiate lower retention in online courses.

Diaz’s (2000) study, discussed in the previous section on comparing online and face-to-face retention, also administered surveys to students. The researcher found online (n = 94) and traditional (n = 40) students had statistically significant differences in learning styles, specifically the independent, collaborative, and dependent scores using the Grasha-Riechmann Student Learning Style Scales (Grasha, 1996). Online students were more inclined to report being independent, so they were also less inclined to report being dependent and less inclined to report having a collaborative learning style. Students also completed satisfaction surveys; results indicated students in both online and face-to-face formats expressed similar satisfaction with the course section (Diaz, 2000). Diaz referenced his own 1999 research which was similar to this dissertation to demonstrate that online students dropped courses more often than traditional students.
Richards and Ridley (1997), in continuing a study mentioned earlier (Ridley et al., 1995), looked at enrollment and persistence factors of online students (n = 69). Students were asked why they chose to enroll in an online course instead of face-to-face; most reasons (86.9%) were related to scheduling conflicts with work, class schedule, or other constraints. Their study suggested student satisfaction with the interface and time or scheduling challenges of face-to-face courses were the main factors causing students to persist in their online courses. Students were also asked to comment on the online delivery system because it had recently progressed from a dial-up modem/bulletin board interface to a web browser interface; 84.6% preferred the new interface over the previous text-based interface. Richards and Ridley (1997) referenced their own previous research (Ridley et al., 1995) in their statement about online students withdrawing more than face-to-face students.

These are a few examples of studies that are commonly conducted to explore online retention. As already mentioned, studies like these usually refer to online retention as being lower than face-to-face retention. Models to explain online retention are the other common area of research in online education.

Models of Online Retention or Persistence

Much of the retention research has been conducted to develop or use models to help explain reasons for student persistence, retention, or attrition. It is important to note that Kelly (1996) in examining retention using Astin’s (1977, 1993) I-E-O model, explained that research showed student characteristics consisted of many factors that did not alone explain retention or attrition. He cautioned against reducing a multi-dimensional event to a one-dimensional comparison.
Distance education and online student retention models will be discussed again in the
next section about models which are relevant to the theoretical framework. Models which are not
included in that discussion are included here as additional examples of studies perpetuating the
idea that online retention is lower than face-to-face retention.

Berge and Huang’s (2004) sustainable student retention model is flexible to fit the
priorities of the institution and includes personal, institutional, and circumstantial variables. The
model allows for prioritizing and connecting variables. Personal variables include demographic
and other personal attributes, such as age, gender, academic skills and abilities, and prior
educational experience. Institutional variables include organizational characteristics and
institutional attitude. Circumstantial variables include perceived stress, responsibilities, and
institutional, social, and academic interactions. Berge and Huang (2004) listed Carr (2000), Diaz
(2002) and Frankola (2001) as researchers who believe online retention is lower than face-to-face
without presenting it as fact.

Angelino and Natvig (2009) proposed a model for engagement of the online learner
based on Tinto’s (1975) student integration model. The model has four strategies: recruitment,
coursework, post coursework, and alumni. Each area has five ways to engage students.
Recruitment engages through marketing, initial contact, potential students, registration, and class
information. The coursework strategic area engages students through a course website, start of
class, assignments, discussions, and course activities. Post coursework consists of pre-
registration, course evaluation, student feedback, celebrate success, and graduation. The alumni
strategy includes alumni association, mentoring, interviews, promote program, and
Sloan Consortium reports and Berge and Huang (2004), saying that attrition is a concern.
Theoretical Framework

A theory “states that things are related in a particular way. A theory is a statement of how things ought to be” (Bouma & Atkinson, 1995, p. 21). Theories or models could be treated as synonyms. If they are treated as different concepts, a theory is considered to be informal while a model is systematic and formal (Thomas, 2006).

Reviewing leading models or theories underlying student retention may help with understanding the different reasons for students taking an online course or a face-to-face course and not completing it. Some current theories or models of retention or persistence in online learning have built upon retention theories from face-to-face instruction. Kember (1989) observed that more attrition research went towards describing attrition rather than developing theoretical models.

Traditional Student Retention Models

Tinto’s (1975) model is used the most often to explain student retention (Simpson, 2003; York, 1999). Building upon Spady’s (1970) and Rootman’s (1972) work and using Durkheim’s theory of suicide as Spady (1970) did, Tinto theorized that an individual’s background and commitments toward completing the goal and to that institution influence the level of social and academic or intellectual integration achieved. This integration continues to influence the commitment levels of goal completion and to the institution. He presumed that either a low goal commitment or institution commitment could be the main factor leading a student to drop out. Tinto (1987, 1993) explained that most individuals choose to leave school based on their perception of how well they integrated socially or intellectually, rather than the reality of their integration. The model acknowledged that “what one thinks is real, has real consequences” (Tinto, 1987, p. 127).
Astin (1977, 1993) presented another widely referenced model which includes retention, or degree completion within four years, along with other college outcomes. Using the input-environment-outcome (I-E-O) model, inputs are the student’s characteristics when entering college. Environment refers to the “programs, policies, faculty, peers, and educational experiences” the student is exposed to. Outcome refers to the student’s characteristics after being exposed to the college environment. Student’s high school grades were the strongest predictor of retention (Astin, 1977, 1993).

**Non-Traditional Student Retention Model**

Bean and Metzner (1985) added to the work of Spady (1970), Tinto (1975), and Pascarella (1980) in developing a model for non-traditional students. This model could be closer to that needed for online learning because it factors in less social integration and more environmental factors. Bean and Metzner (1985) defined non-traditional students as older, part-time, or living off campus, or any combination of the three.

The model consists of four main groups of variables:

- **Background and defining variables** consisting of (a) age, (b) enrollment status, (c) residence, (d) educational goals, (e) high school performance, (f) ethnicity, and (g) gender.
- **Academic variables** include (a) study habits, (b) academic advising, (c) absenteeism, (d) major uncertainty, and (e) course availability.
- **Environmental variables** are (a) financing, (b) hours of employment (c) outside encouragement (d) family responsibilities, and (e) opportunity to transfer.
- **Psychological outcomes** include (a) utility, (b) satisfaction, (c) goal commitment, and (d) stress.
The non-traditional student does not usually change his or her social group in order to attend college, therefore social interaction is viewed as less of a factor than in Tinto’s (1975) model. Bean & Metzner (1985) considered environmental variables to be more important than academic variables to nontraditional students. When both variables were good, students remained in school; when both variables were bad, the students left school. Students were also likely to depart school if environmental variables were bad, even if academic variables were good. The student was likely to remain in school when academic variables were bad if environmental variables were good.

*Distance Education Retention Model*

Kember (1989) developed a longitudinal model or framework for distance education student drop outs by starting with the Tinto (1975) model, which had the most empirical research to support it; it also had some applications in distance education. Based on research, Kember (1989) added more student background variables, such as home and work. Instead of including both goal and institutional commitment, the model divided goal commitment into intrinsic motivation and extrinsic motivation. The model adds work as part of the social environment and integration aspects of the model, since for distance students, both academic environment and work/social environment are different than for face-to-face students. Finally, a cost/benefit analysis precedes the dropout decision.

At the time Kember (1989) developed the model, distance learning was discussed in terms of having telephone conferences and mailing course content packages. Most courses were delivered through print media, but audio and video cassettes and computer assisted learning were mentioned as alternatives.
Online Student Retention Model

Rovai (2003) developed a *composite persistence model*, shown in Figure 2.1, based on the models by Tinto (1975, 1987, 1993) and Bean and Metzner (1985). He combined elements of the two models with elements of other research to reflect needs of online students, such as skills and other special needs for online or distance students, and a need to integrate learning styles of online students with teaching styles. The resulting model reflects factors before and after students are admitted.

Before admission, the factors include:

- Student characteristics of age, ethnicity, gender, intellectual development, academic performance, and academic preparation (from Bean and Metzner (1985) and Tinto (1975)).
- Student skills of computer literacy, information literacy, time management, reading and writing, and computer-based interaction.

After admission, internal factors are viewed as being more important than external factors, such as:

- From Tinto (1975): academic and social integration, goal and institutional commitment, and learning community.
- From Bean and Metzner (1985): study habits, advising, absenteeism, course availability, program fit, current GPA, utility, stress, satisfaction, and commitment.
- Student needs: clarity of programs, self-esteem, identification with school, interpersonal relationships, and accessibility to services.
- Pedagogy: learning styles and teaching styles.

External factors included finances, work hours, family responsibilities, outside encouragement, opportunity to transfer, and life crises (from Bean and Metzner (1985)). If internal factors are
weak, then external factors are a larger factor in a student’s persistence; students are more likely
to withdraw if external factors are weak as well.

**Summary**

This chapter reviewed the literature comparing online retention with face-to-face retention over the past ten years. Eleven studies were primarily conducted at community colleges or individual course offerings at universities, indicating a need for studies on a more encompassing scale. Most authors appeared to consider online retention as lower than face-to-face retention. Some authors recognized the discrepancy or debate in comparing online and face-to-face attrition or retention.
The theoretical framework was discussed in this chapter and was used to help explain results in chapter five.
CHAPTER 3
METHODOLOGY

The overarching purpose of the study was exploratory in nature and intended to help describe the current state of the differences between online and face-to-face course completion rates. Exploratory research is intended to address conditions in which there has been little empirical research in a topic, the topic has matured, and there has been enough change in the topic that it needs explored again (Stebbins, 2001).

The purpose of this chapter is to explain the methodology which led to the methods used in the study. The methodology was primarily influenced by the research questions and the literature review and was revisited and refined throughout the study. During development of the methodology, the focus was on answering the research questions. While determining the research design and the methods, the nature of the data needed to answer the research questions guided decisions, and the focus remained on answering the research questions.

Paradigm or World View

There are different viewpoints on which comes first and are directly relevant to the study reported here: the research question or the theoretical perspective, paradigm, or worldview. Crotty (1998) stated that we first select our method and methodology to answer the research questions and then decide on the theoretical perspective and epistemology, but presented several diagrams which showed the epistemology informing the perspective which then informs the methodology and the method, as if they could be reversed. Other authors explained that the research questions and methods are influenced by the researcher’s worldview or theoretical perspective (Clark & Creswell, 2008), either consciously or subconsciously (Hesse-Biber, 2010).
Clark and Creswell’s (2008) discussion of paradigms and the way in which Thomas Kuhn used the term may help explain some of the differences of opinion. They observed that Kuhn used the word paradigm in at least four basic ways. According to their chart, paradigm could be a worldview, an epistemological stance, shared beliefs, or a model example (Clark & Creswell, 2008). Masterman (1970) determined Kuhn used paradigm in over 20 different ways and went on to list 21 of them. Time may explain the differences as well; writers may change their mind on terminology usage, but the book or journal they wrote before is still widely read. Clark and Creswell (2008) wrote that Kuhn “wished that he had used a different term like disciplinary matrix” (p. 32).

Kuhn (1970) wrote a post-script to the second edition of The Structure of Scientific Revolution, explaining that he corrected errors for the second edition but did not change the overall content. He did address the use of the word paradigm, explaining that he used the term in two different senses. The first sense referred to “the entire constellation of beliefs, values, techniques, and so on shared by the members of a given community” and the second to “one sort of element in that constellation, the concrete puzzle-solutions which, employed as models or examples, can replace explicit rules as a basis for the solution of the remaining puzzles of normal science” (Kuhn, 1970, p. 175).

The postpositivist worldview or paradigm is appropriate for a study to answer quantitative questions (Creswell, 2014). One view of postpositivism is that “knowledge is produced through evaluation of evidence and causal events; it cannot, however, be absolutely secure” (Paul, Graffam, & Fowler, 2005, p. 46). The three authors considered postpositivism similar to common sense in that humans interact within the world to achieve a purpose, inquiring
and acquiring knowledge that is acted upon until they acquire new and better knowledge to replace it.

Gall, et al. (2007) explained six characteristics of postpositivism which make it appropriate for educational research. Postpositivism allows for created and shared concepts and procedures, replicable findings, refutable claims of knowledge, bounds on claims of knowledge, error and bias control, and a moral commitment to discourse which allows for progress.

Research Design

The quantitative analysis was a non-experimental, causal-comparative design which can be used to help determine results of a difference, a cause, or an independent variable that exists between or among groups and has already happened (Creswell, 2014; Fraenkel et al., 2011). This causal-comparative design was not intended to identify a cause for any difference; it was only intended to indicate if a difference existed. Causal-comparative studies can identify a need that leads to other research (Fraenkel et al., 2011).

Fraenkel et al. (2011) explain that causal-comparative research differs from correlational research in that the former compares two or more groups, while the latter compares individuals’ scores on each variable. Usually causal-comparative studies analyze at least one categorical variable (group membership) while correlational studies analyze two or more quantitative variables. Finally, correlational studies usually produce scatterplots and correlational coefficients while causal-comparative studies produce comparisons of averages and crossbreak tables. The quantitative analysis is explained in more detail in the next sections.

Research Questions

The research questions which were best addressed using a quantitative design were:
1. What is the difference among select four-year universities’ course completion rates of online distance learning courses across the regions of the United States?

2. What is the difference in online distance learning course completion and face-to-face course completion among select four-year universities across the regions of the United States?

3. What is the difference in online distance learning course completion between undergraduate and graduate students at select four-year universities across the regions of the United States?

Population Sample

A purposive sample of course completion data from select large four-year public universities was collected for a comparative analysis. A purposive sample may be used if needed specifically for the research and if researchers have knowledge of the population (Fraenkel, Wallen, & Hyun, 2011). Extending the study beyond one university helps avoid “misinformed educational policies” (Ma, Ma, & Bradley, 2008, p. 63).

Groups should be homogenous on some key variables, or it will be likely no differences will be found between groups (Fraenkel et al., 2011), therefore a form of geographical cluster sampling was used to select comparable universities nation-wide. Four universities were first selected in one state based on their operation as universities rather than being part of other university systems. Nationwide, other universities were chosen to parallel the four original universities by Carnegie Classification – large and high or very high research. Most requests were dispersed among different states in the regions; at least four universities were asked to represent each region. Selecting only large, high or very high research, four-year universities provided for a more homogenous population than selecting a larger variety of colleges and universities, which is important for a comparison study (Fraenkel et al., 2011).
As discussed, at least four universities from each region of the United States as established in IPEDS ("College & Careers," n.d.) were asked to participate in the study. The regions and associated states are:

1. New England (CT, ME, MA, NH, RI, and VT)
2. Mid East (DE, DC, MD, NJ, NY, and PA)
3. Great Lakes (IL, IN, MI, OH, and WI)
4. Plains (IA, KS, MN, MO, NE, ND, and SD)
5. Southeast (AL, AR, FL, GA, KY, LA, MS, NC, SC, TN, VA, and WV)
6. Southwest (AZ, NM, OK, and TX)
7. Rocky Mountains (CO, ID, MT, UT, and WY)
8. Far West (AK, CA, HI, NV, OR, and WA)

In IPEDS there is an Outlying Areas region which was not invited to be included in the study (AS, FM, GU, MH, MP, PR, PW, and VI) because there is no large public university in those locations.

The specific Carnegie classification variables which were used to help select comparable universities were: control (public), size and setting (large: 15, 16, and 17), and basic 2010 (high research or very high research: 15 and 16). The Carnegie Foundation website ("Institution Lookup," n.d.) provided a search method which allowed for displaying similar institutions but started moving to university control during this study and made a spreadsheet available for download instead.

**Data Collection**

Thirty-three selected universities were asked to provide data on course completion of both online and face-to-face courses for academic year 2013-2014. The researcher contacted the
universities using electronic mail or a request form on the university website. Some universities received follow up telephone calls. Data collection took place in February and March 2015.

The contacted universities were distributed throughout IPEDS Regions 1-8, with five universities from Region 1 and four universities each from Regions 2-8. A sample electronic mail is at Appendix A. The researcher sent follow-up requests to universities approximately every two weeks until follow-ups had exhausted the responses. Of the universities contacted, eleven universities provided data, thirteen declined to participate, three failed to respond, and six responded that they may provide data before the study concluded.

Universities are anonymized. Throughout the study, universities are referred to by their Integrated Postsecondary Education Data System (IPEDS) region followed by an alpha character. Region 6, for example, had four universities participate by providing data, so the identifying codes range from 6A to 6D. Only Region 4 through Region 8 participated in the study.

The university offices of institutional research provided enrollment and course completion data for the study by completing a spreadsheet the researcher provided (at Appendix B). The institutional research offices at the individual universities provided aggregated data directly to the researcher. Initially, state-level data repositories were contacted to see if they had the enrollment and course completion information necessary. Seven states were contacted and one state responded with data. This data was not used in the study because it was not broken out by classification level and three universities in the state had already provided data. At the conclusion of the study, participating universities were provided the study and told which university they represent in the study; one university charged the researcher for two hours of work. No other incentives were offered.
Variables

The independent variables were course delivery method (categorical: online or face-to-face) and student classification (categorical: undergraduate or graduate). The dependent variable was course completion rate (continuous: percentage of students) or actual enrollment and course completion numbers (categorical: course completed) when using numerators and denominators instead of percentage. Two other independent variables, region and university, were available to help answer the research questions.

Eleven universities provided enrollment information for both fall and spring semesters separately for academic year 2013-2014. The quantity of students was duplicated for students taking more than one course and for more than one semester; one university reported removing duplicate students. The university offices of institutional research provided enrollment and course completion data for the study by completing a spreadsheet the researcher provided (at Appendix B).

Data Analysis

Analysis was conducted using more than one software program. Microsoft Excel was used to collect, transform, and compile data and calculate ratios. Descriptive and inferential statistics and were prepared using SPSS 22.0. A website (Eck & Ryan, n.d.) was used to calculate statistical significance of a chi-square goodness of fit. A downloadable spreadsheet (DeCoster, 2012) was used to convert effect sizes.

Analysis of continuous variables, course completion rates or percentages, was originally planned using a chi-square test for independence for RQ1 and analysis of variance (ANOVA) for the means of the percentages to answer RQ2 and RQ3. When results from the ANOVA calculated the means differently for semester and student classification than what had been
calculated in Microsoft Excel, it became apparent that percentages should not be analyzed directly. Further research showed that chi-square analysis could be used to answer all three research questions by using course completed and not completed numbers in a contingency table. Enrollment numbers needed to be weighted within SPSS on frequency to keep SPSS from analyzing the numbers as continuous variables and instead treat the numbers as a categorical variables. The issues with percentages are explained in more detail starting in the next section.

Quantitative Analysis Using Ratios and Aggregate Data

Analysis of ratios or proportions in aggregate data can be considered “spurious” (Pearson, Lee, & Elderton, 1910, p. 534) due to other factors or variables beyond the ratio (Firebaugh & Gibbs, 1985). Hox (2010) explained that using a proportion as a dependent variable violates assumptions of continuous scores, normality and homoscedasticity. Robinson (1950/2009) explained that correlations drawn from aggregate data often do not match correlations drawn from individual data.

Ratios or Proportions

The example Pearson et al. (1910) used to illustrate an issue with proportions was a comparison of death rates between different populations—the age distribution of the populations should be a variable considered along with the death rate. Firebaugh and Gibbs (1985) explained that the issues may be spontaneity bias, or “bias due to feedback effects” and use of the wrong variables in the model (p. 713).

Edward Simpson observed a similar effect which was later called Simpson’s paradox (Pearl, 2014). Pearl (2014) explains that Simpson’s paradox occurs when a combined table shows the reverse rate of a disaggregated table, or the average of a product does not equal the product of the averages. Wainer and Brown (2011) demonstrated how this appeared when
comparing math scores by race from one state to another, as shown in Table 3.1. State scores overall were higher in Nebraska, but when viewed by race, New Jersey had higher scores.

Table 3.1.

*Example Demonstrating Simpson’s Paradox (Wainer & Brown, 2011, p. 899)*

**NAEP 1992 8th grade Math Scores**

<table>
<thead>
<tr>
<th>State</th>
<th>White</th>
<th>Black</th>
<th>Non white</th>
<th>Standardized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nebraska</td>
<td>277</td>
<td>281</td>
<td>236</td>
<td>259</td>
</tr>
<tr>
<td>New Jersey</td>
<td>271</td>
<td>283</td>
<td>242</td>
<td>260</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Proportion of population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nebraska</td>
</tr>
<tr>
<td>New Jersey</td>
</tr>
<tr>
<td>Nation</td>
</tr>
</tbody>
</table>

Wainer and Brown (2011) created a standardized score based on the national demographic mix, multiplying the average score for the subgroup by the respective national demographic percentage. New Jersey’s standardized score is calculated as follows: 

\[(283 \times 0.69) + (242 \times 0.16) + (260 \times 0.15) = 273.\]

Zuzovsky, Steinberg, and Libman (2011) explained that Simpson’s paradox occurs when subgroups are not distributed equally at the different levels in the relevant contextual variable. Achievement for the population and the relationship between the contextual variable and the achievement can both be reported in a misleading manner. The authors used analysis of
covariance to control for the effect and applied standardization as Wainer and Brown (2011) recommended.

Standardization is a type of weighted average which takes population composition into account and helps researchers compare across groups (Schoenbach & Rosamund, 2000). Scores or percentages can be standardized or adjusted using various methods. Indirect standardization is the most common, followed by direct standardization. Both methods include the crude rate of the standard population in the calculation. Indirect standardization uses rates from the study population; direct standardization uses weights from a standard population. There are advantages and disadvantages to the use of standardized rates. Both indirect and direct standardized rates only have meaning when compared with a rate adjusted using a similar method. Opponents of standardization may recommend analysis of at least two comparison summaries when determining differences (Fleiss, Levin, & Paik, 2003).

Snedecor and Cochran (1989) discussed three other approaches to analyzing proportions: weighting, arcsine transformation, and logit transformation. Warton and Hui (2011) considered arcsine analysis outdated and no longer necessary to use since the advent of computer-based statistical analysis. Snedecor and Cochran (1989) explained that arcsine transformation does not adjust for variance inequalities caused by having different denominators and advise use of a weighted analysis if denominators have a wide variance.

Logit is another common transformation used for proportions (Hox, 2010). Warton and Hui (2011) proposed use of the logit transformation instead of arcsine because it correctly models and maps to the original line while arcsine approximates it. Logit is the inverse of the logistics function and is considered to be relatively easy to calculate (Cramer, 2003).
Aggregate or Group Data

Firebaugh (1980) compared two regression methods most often recommended for nonexperimental assessment of group effects: covariance analysis and contextual analysis. He concluded that covariance analysis best addresses how large the total impact of the group is and contextual analysis, which some researchers call multilevel analysis (Roux, 2002), or hierarchical linear modeling (Raudenbush & Bryk, 2002), helps identify the effect group characteristics have. Both methods can be used, as “covariance analysis may be useful for assessing the adequacy of the contextual model” (Firebaugh, 1980, p. 21).

Heck, Thomas, and Tabata (2014) explained that not using a multilevel model can lead to findings which show a higher significance level than if clustering were considered. Standard errors may be underestimated without a multilevel model, which calculates standard errors based on a ratio of the estimate. Bowman (2012) recommended multilevel analysis when conducting a quantitative meta-analysis for the same reason multilevel analysis is recommended for other studies: independence of observations are often violated. Bowman (2012) uses the specific example of having “students at various institutions within the same sample” (p. 378).

Multilevel analysis can prevent misleading conclusions which may result from analyzing data that is actually at more than one level. Heck et al. (2014) stated that “because statistical tests of model parameters are based on the ratio of an estimate to its standard error, the underestimation of standard errors will often lead to more findings of significance than would be observed if clustering were considered” (p. 7). Clustered data such as from individuals in the same geographic area are likely to be correlated; not taking this into account during analysis could lead to misleading results, for example the standard deviation of the mean difference in weights between groups could be overestimated by assuming independence (Dobson, 2002).
Multilevel analysis may be conducted without violating assumptions of independence by using random effects for intercepting variables in the model (Tabachnick & Fidell, 2013). It is more suitable for group level analysis than other models, such as analysis of variance and multiple regression, because it can examine correlated data and unequal variances within groups. Group level variability and individual level variability are both maintained during multilevel analysis (Heck, Thomas, & Tabata, 2012).

Heck et al. (2014) explained that weighting is not a feature with multilevel models in SPSS; analysis requiring use of weights should be done with another software program, or the researcher should revert to a single-level model. If weights are only at level two, or where the clusters were sampled with unequal probability, a method for single level weighted analysis can be used (Asparouhov, 2006).

Generalized linear mixed modeling (GLMM) is an extension of multilevel modeling which can be used to analyze categorical variables (Heck et al., 2012). This method of data analysis was late to develop in SPSS because of the additional complication in solving math problems for nonlinear equations; categorical responses do not usually follow the normal distribution. A dichotomous outcome, for example, reflects one of two responses for an individual, such as passing or failing a course (Heck et al., 2012). “Using ANOVA or ‘general’ linear model methods does not exempt you from GLMM-related issues—it simply makes you oblivious to their existence” (Stroup, 2013, p. xvii).

Transformation (Heck et al., 2012) is also often used when data is not normally distributed or contains heteroscedastic errors, meaning there are problems related to different variances. Transformation is often not appropriate when the variables are dichotomous, as in GLMM. A dichotomous variable is usually coded as 1 or 0, so the transformation will result in a
0 or produce an error. Count data would also most likely violate assumptions of normality, especially in the case of students passing a course. Under GLMM, transformations and error distribution decisions are handled directly by the model. Heck et al. (2012) recommended selection of the robust standard error option to adjust for mild violations of the normal distribution (Heck et al., 2012).

As a multilevel analysis using course completion rates, centering is not required if raw data is not being analyzed. Also, course completion is the dependent variable; centering is more commonly used with independent variables (Enders & Tofighi, 2007). Centering could be considered if using raw course enrollment numbers as an independent variable (Tabachnick & Fidell, 2013).

Before running multilevel analysis, the null model with no predictors should be examined for both within group variance and between group variance. If the intraclass correlation (ICC), which is a proportion of the variance between groups compared to the total variance, is smaller than .05, there would not be much benefit to proceeding with the multilevel analysis (Heck et al., 2014; Raudenbush & Bryk, 2002).

Model fit within SPSS can be determined by selecting models with a lower Akaike’s information criteria (AIC) or Bayesian information criteria (BIC) (Heck et al., 2012). In a comparison of AIC and BIC, BIC was judged to be more consistent and better suited for selecting the true model; AIC is better suited for predicting future data (Kuha, 2004). Multilevel analysis for this study more closely resembled an explanatory approach rather than a predictive approach (Heck et al., 2014).
Effect Size

Tabachnick and Fidell (2013) considered effect size in multilevel linear modeling to be “less straightforward” than with other methods of analysis (p. 790). They explain that ICC (ρ) should only be calculated for the null model or there would be “different correlations for cases with different values of a predictor” (p. 826) and recommended an $R^2$ effect size calculation offered by Kreft and De Leeuw (1998). The estimate does not fit most models, as it should only be used on models with random intercepts and not on predictors with random slopes; separate calculations are needed for level one (within group) and level two (between group) (Tabachnick & Fidell, 2013).

Nakagawa and Schielzeth (2013) recommended the use of $R^2$ for HLMs and GLMMs because it is “unitless” (p. 133) and can be compared across studies as are other standard effect size measures. The authors presented both marginal and conditional $R^2$ formulas. Marginal $R^2$ formulas concern variances explained by fixed factors, and conditional $R^2$ formulas concern variances explained by both fixed and random factors. They also recommended separate formulas for similar and dissimilar cluster sizes. Nakagawa and Schielzeth (2013) surmise that the $R^2$ formulas presented for both HLMs and GLMMs are less subject to implementation problems than other $R^2$ formulas.

Feingold (2009) explained that part of the challenge is there are different ways effect size has been calculated with growth modeling analysis (which includes multilevel analysis), either for power calculations or for magnitude effect. Power calculations use the standard deviation of changed scores while magnitude effect uses the raw scores.

Hedges (2007) recommended selecting effect size computations based on the intended studies with which they will be compared. Computations should also be selected based on
whether cluster sample sizes are equal or unequal. Hedges (2007) divided effect size for cluster-randomized designs into three types: within cluster, within treatment, or between clusters.

Snijders and Bosker (2012) explained that ICC is appropriate to use as an effect size indicator in multilevel analysis since it is transforming both the within group and the between group variances. In interpreting the ICC (ρ) as an effect size, Tymms (2004) recommended the following formula: \( \Delta = [4\rho/(1- \rho)]^{1/2} \) (p. 65).

Spybrook (2008) recommended a standardized effect size which does not depend upon the scale of the outcome variable. The estimated standardized effect size, \( \hat{\delta} \), is the difference between the group sample means divided by the estimated standard deviation of the outcome variable: \( \hat{\delta} = (\bar{Y}_E - \bar{Y}_C) / [(\sigma^2 + \tau)^{1/2}] \) (Spybrook, 2008, p. 285). The Cohen (1988) effect size interpretation suggestions of .80 and above for large, between .80 and .50 for medium, and between .50 and .20 for small can be used. Spybrook (2008) cautioned that effect sizes smaller than .20 in education should not necessarily be ignored.

For GLMM, Heck et al. (2012) suggested converting the log odds to an odds ratio (\( e^\beta \), where \( e \) is Euler’s number, about 2.71828 and \( \beta \) is the log odds). Borenstein, Hedges, Higgins, and Rothstein (2009) supported the conversion of an odds ratio used for binary data to Cohen’s \( d \) when studies are comparable to each other using \( d = \text{LogOddsRatio} \times [(3^2) / \pi] \) (p. 47) and also present a conversion from log odds ratio to odds ratio as \( \text{LogOddsRatio} = \ln (\text{OddsRatio}) \) or \( \text{OddsRatio} = \exp (\text{LogOddsRatio}) \) (p. 36).

Contingency Table

A cross-tab procedure or contingency table was created in SPSS in order to conduct a chi-square test for independence; this test allowed for analysis of course enrollment numbers instead of course completion percentages. Chi-square is appropriate when analyzing two
categorical or nominal variables. The data should represent a large sample size or a fairly even spread of the subjects among the levels or categories (Morgan, Leech, & Gloeckner, 2013).

Data was reformatted from long to wide and weighted on frequency of course completion status in SPSS. Analysis was performed using course completion status (completed or not completed) as the dependent variable for RQ1; data was first filtered to show online enrollments only. In exploring RQ2, course completion status (completed or not completed) by course mode (online or face-to-face) was selected. For RQ3, student classification (undergraduate or graduate) by completion status (completed or not completed) was analyzed after filtering for online course enrollment only.

Statistical significance is based on the $p$ value and indicates the conditional probability of the data assuming the null hypothesis is true, with the null hypothesis being that there is no difference. Other conditions usually need to be met as well, such as random sampling, independent and reliable scores, and distribution requirements of normality and homoscedasticity (Kline, 2013). Assumptions are different for the chi-square test for independence, which is a non-parametric test. Assumptions include independence of individual observations (no observation has an effect on another observation) and requirements for expected cell frequencies size – preferably five or greater (Weinberg & Abramowitz, 2008).

A chi-square test for independence was conducted between universities and online course completion for research question 1. For research question 2, a chi-square test for independence was conducted between universities and both online and face-to-face course completion. Research question 3 consisted of using the chi-square test for independence between universities, online distance learning course completion, and student classification (undergraduate or graduate).
The cross-tab was created both with the university variable and without for all three quantitative research questions. All cell frequencies were greater than five. Face-to-face mode was filtered out of the analysis for research questions 1 and 3. A chi-square goodness of fit test with 1 degree of freedom was used to determine the significance of the differences between the dependent variable’s observed and expected values; continuity correction was not applied. The chi-square for independence procedure calculated the expected and actual rates based on all online course completion data for all universities in the study. The number of students who did not complete courses was entered along with the number of students who did complete online courses. The sum of each total was used as the basis for calculating an expected number of students to complete and not complete online courses.

Snedecor and Cochran (1989) show an alternative for computing chi-square, or the variance test for homogeneity of the binomial distribution, which can help investigate if “true proportions differ from sample to sample” (p. 202). The formula includes both the denominator and the proportion and is considered quicker to compute than other chi-square calculations: 

\[ X^2 = \sum n_i (p_i - \bar{p})^2 / (\bar{p} x \bar{q}) \]

The \( X^2 \) calculation using the formula in MS Excel that Snedecor and Cochran (1989) demonstrated yielded the same result in SPSS.

**Multilevel Analysis**

The issue with ratio analysis which was discussed earlier in this chapter became apparent when the ratios calculated in SPSS for each university case or row for online undergraduate and graduate students did not equal the total ratio already calculated in a separate spreadsheet. Standardization of scores, similar to Wainer and Brown (2011)’s approach, was considered but the requirement for a weight or a rate from a standard population was a restriction. As discussed in the introduction, a national level rate for online course completion does not currently exist.
The indirect standardized method would have been the method used in order to reflect the study population’s weight distribution of student enrollment.

Other alternatives were considered to transform the data for analysis. The arcsine was produced in Microsoft Excel using the following formula: \( \text{ASIN}(\sqrt{\frac{\text{Numerator}}{\text{Denominator}}}) \). The logit was produced with this formula: \( \text{LOG}((\frac{\text{Numerator}}{\text{Denominator}})/(1-(\frac{\text{Numerator}}{\text{Denominator}}))) \). Analysis could have proceeded following the conversion and then the results could have been converted back to a proportion (Stroup, 2013). An SPSS analysis to compare the results of the conversions is presented in Appendix C for informational purposes only. Multilevel modeling was also attempted using percentages, but an extension of multilevel modeling, GLMM, appeared more suited for the data analysis.

Generalized linear mixed modeling (GLMM) was conducted in SPSS rather than analyzing logit or other transformed variables. Since the study consisted of aggregate data at two levels (semester and university level), as well as percentages, multilevel analysis and meta-analysis were also considered. The GLMM procedure accounts for challenges presented by both group-level data and percentages (Heck et al., 2012).

The dependent variable and independent variables were included in the model with a view towards determining magnitudes of effect. Heck et al. (2012) advised using theory to build the model and starting first with fixed slope effects at level 1. The authors suggested that additional random slopes can be added later as needed, still keeping the theoretical focus and research purpose in mind.

An unconditional or null GLLM analysis with no predictors or fixed effects was conducted in SPSS to help determine if level 2, university, impacted the analysis (Heck et al., 2012). Within the SPSS mixed function for GLMM, the numerator (number of courses
completed) and denominator (number of courses conducted) were input and the logit link function was selected. A statistically significant random intercept variance across groups indicated the analysis was warranted; this will be shown in chapter 4.

In order to address research question 1 and 3, a model with binomial probability distribution and logit link function using robust estimation was constructed in SPSS with the random effect of university at level 2; research data was filtered to show only online students. The ICC was calculated using the results of the model: \( \rho = \frac{\sigma^2_{\text{Between}}}{\sigma^2_{\text{Between}} + 3.29_{\text{Within}}} \). The use of 3.29 represents the variance of a logistic distribution with a 1.0 scale factor, calculated as \( \pi^2/3 = 3.29 \) (Heck et al., 2012; Hox, 2010). The fixed effect and random effect of student classification was present in the final model after adjusting until a balance was achieved between obtaining the lowest BIC level while still keeping within the context of the research questions (Heck et al., 2012; Seltman, 2014).

In order to address research question 2, a similar model with binomial probability distribution and logit link function using robust estimation was constructed in SPSS with the random effect of university at level 2; the filter in SPSS was removed to show both face-to-face and online students. The ICC was calculated using the results of the model: \( \rho = \frac{\sigma^2_{\text{Between}}}{\sigma^2_{\text{Between}} + 3.29_{\text{Within}}} \). The fixed effects and random effects of student classification and course mode were both present in the final model after adjusting until a balance was achieved between obtaining the lowest BIC level while still keeping within the context of the research question (Heck et al., 2012; Seltman, 2014).

Effect sizes were calculated using the log odds ratio from SPSS; a previously discussed formula and a website (DeCoster, 2012) were used to convert the log odds ratio to an odds ratio and Cohen’s \( d \).
Reliability and Validity

Reliability “refers to the consistency of the scores obtained” (Fraenkel et al., 2011, p. 154). In the case of secondary data, which was originally obtained for different purposes (Smith, 2008), reliability will depend on the provider of the data. Smith (2008) cautioned that secondary data is like other data and probably contains errors.

Validity refers to the “appropriateness, correctness, meaningfulness, and usefulness of the specific inferences researchers make based on the data they collect” (Fraenkel et al., 2011, p. 148). A local expert was asked to review and validate the definitions and requests for data that were sent to institutions. Missing data is an issue with multilevel modeling, as with most statistical analysis, but for this study all missing data was addressed by each university. There was no missing data to make decisions on in the aggregate analysis. Threats to internal validity of causal-comparative research include inability to randomly assign subjects to the groups and inability to manipulate the independent variable (Fraenkel et al., 2011). There is a likelihood the groups which participate in face-to-face learning and online learning are not equivalent on important variables (Distance Education and Training Council, 2007).

Ethical Issues

Ethical propriety was maintained throughout the research process. An institutional review board (IRB) application was filed by the major professor through the university in December. Data collection began in February 2015. Anonymity is being kept for the universities included in the study. All forms will be destroyed at a later date in accordance with university policy.

Summary

This chapter has explained the research design and the methods used in the quantitative analysis. The analysis was complicated by data which used percentages and the ramifications of
analyzing aggregate group-level data. Simpson’s paradox can be problematic with percentage data. Aggregate level data analysis may not account for group effects. Chi-square analysis was first conducted, but it did not account for group effects. After careful consideration, generalized linear multilevel modeling was approached as a remedy for percentages and aggregate group-level data. The next two chapters report and then discuss the results.
CHAPTER 4
RESULTS

The purpose of the study was to compare retention or course completion in online learning and face-to-face learning at select four-year public universities. A causal-comparative design was chosen to determine if statistically significant differences existed between course mode, student classification, and universities. Statistical significance is based on the $p$ value and indicates the conditional probability of the data assuming the null hypothesis is true, with the null hypothesis being that there is no difference. The quantitative analysis looked for differences in course completion rates among participating universities. Analysis to answer the three research questions was first conducted using a chi-square test for independence and then conducted using generalized linear multilevel modeling (GLMM). Both methods allowed for analysis of the course enrollment numbers instead of the course completion rates or percentages.

Research Questions

Results are in order by research question. The research questions were:

1. What is the difference among select four-year universities’ course completion rates of online distance learning courses across the regions of the United States?

2. What is the difference in online distance learning course completion and face-to-face course completion among select four-year universities across the regions of the United States?

3. What is the difference in online distance learning course completion between undergraduate and graduate students at select four-year universities across the regions of the United States?

Descriptive Statistics

The face-to-face course enrollment ($n = 2,288,007$) average completion rate was 92.6% (SD = 4.25). The online course enrollment ($n = 237,499$) average completion rate was 89.7 (SD=
Table 4.1 shows the mean overall course completion rates, including undergraduate and graduate. Undergraduate online enrollment was the only rate below 90%.

Table 4.1.

*Mean Overall Course Completion Rate for Fall 2013 and Spring 2014*

<table>
<thead>
<tr>
<th>Course Format</th>
<th>M</th>
<th>SD</th>
<th>Undergraduate</th>
<th>Graduate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online</td>
<td>89.71</td>
<td>5.95</td>
<td>86.37</td>
<td>93.06</td>
</tr>
<tr>
<td>Face-to-face</td>
<td>92.58</td>
<td>4.25</td>
<td>90.42</td>
<td>94.73</td>
</tr>
</tbody>
</table>

Overall completion percentage rates were normally distributed, D (88) = .065, p = .200. Online percentage rates appeared to not be normally distributed, W (44) = .925, p = .007. Further analysis indicated graduate student online completion rates were normally distributed, W (22) = .962, p = .537; undergraduate student online completion rates appeared to not be normally distributed, W (22) = .900, p = .030, and were they negatively skewed (z-score = -2.41). The Kolmogorov-Smirnov (D) test for normal distribution may be used for samples greater than 50 and the Shapiro-Wilk (W) test for normal distribution may be used for samples less than 50 (Mayers, 2013).

Most universities counted students for each course the students enrolled in. Students often would have enrolled in multiple courses. Therefore, in this study, n does not equal the actual number of students, but represents the number of courses a student was enrolled in, completed, or did not complete. Table 4.2 presents the descriptive statistics for the total course enrollments. As shown, the number of face-to-face students was greater than the number of online students, just as the number of undergraduate students was great than the number of graduate students.
Table 4.2.

*Descriptive Statistics of Overall University Enrollments for Fall 2013 and Spring 2014*

<table>
<thead>
<tr>
<th>Enrollment</th>
<th>Under Graduate</th>
<th>Graduate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>54,072.09</td>
<td>8,972.91</td>
</tr>
<tr>
<td>SD</td>
<td>57,850.03</td>
<td>10,949.69</td>
</tr>
<tr>
<td>n</td>
<td>2,379,172</td>
<td>394,808</td>
</tr>
<tr>
<td>Online</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>10,605.95</td>
<td>1,744.45</td>
</tr>
<tr>
<td>SD</td>
<td>12,411.43</td>
<td>1,487.34</td>
</tr>
<tr>
<td>n</td>
<td>233,331</td>
<td>38,378</td>
</tr>
<tr>
<td>Face-to-face</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>97,538.23</td>
<td>16,207.36</td>
</tr>
<tr>
<td>SD</td>
<td>52,345.26</td>
<td>11,567.68</td>
</tr>
<tr>
<td>n</td>
<td>2,145,841</td>
<td>356,430</td>
</tr>
</tbody>
</table>

Across the individual universities, course completion rates appeared normally distributed for 4A, 5A, 6B, 6C, 7A, 8A, and 8C, but not for 5B, 6A, 6D, and 8C, as indicated by the Shapiro-Wilk test for normality in Appendix D. This test using percentages and not course completion numbers examined whether the course completion percentage was significantly different from a normal distribution for each university. Each university had eight data points—four for each semester (fall and spring) which were divided into course mode (on line and face-to-face) student classification (undergraduate and graduate). The test of for normal distribution
was not a requirement for data analysis but is provided for information. As discussed in chapter 3, proportions are not usually normally distributed and this is one of the challenges of analyzing proportions directly.

Table 4.3

*University Enrollment and Completion Rates*

<table>
<thead>
<tr>
<th>University</th>
<th>Online Course Completion Rates</th>
<th>Face-to-Face Course Completion Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>n</em></td>
<td>%</td>
</tr>
<tr>
<td>4A</td>
<td>7,265</td>
<td>87.5</td>
</tr>
<tr>
<td>5A</td>
<td>11,031</td>
<td>85.6</td>
</tr>
<tr>
<td>5B</td>
<td>104,035</td>
<td>87.6</td>
</tr>
<tr>
<td>6A</td>
<td>30,468</td>
<td>88.5</td>
</tr>
<tr>
<td>6B</td>
<td>34,001</td>
<td>83.7</td>
</tr>
<tr>
<td>6C</td>
<td>22,865</td>
<td>94.1</td>
</tr>
<tr>
<td>6D</td>
<td>6,424</td>
<td>93.3</td>
</tr>
<tr>
<td>7A</td>
<td>20,561</td>
<td>88.7</td>
</tr>
<tr>
<td>8A</td>
<td>16,235</td>
<td>87.3</td>
</tr>
<tr>
<td>8B</td>
<td>7,514</td>
<td>92.7</td>
</tr>
<tr>
<td>8C</td>
<td>11,310</td>
<td>73.0</td>
</tr>
</tbody>
</table>

All university enrollment numbers and completion rates are summarized in Table 4.3. Appendix E shows all enrollment and course completion information by region, university, and semester and is an illustration of the eight data points for each university. Figure 4.1 shows the frequencies of enrollment and course completion for online students. This graphically shows the range for course enrollment numbers, from the lowest enrollment of 6,424 to the highest enrollment of 104,035. Figure 4.2 follows, which shows the frequencies of enrollment and
course completion for face-to-face students. The range for face-to-face enrollments was from 49,740 to 419,983.

Figure 4.1 Frequencies of the online student enrollments and courses completed.

Figure 4.2 Frequencies of the face-to-face student enrollment and courses completed.

Non-parametric statistics in the form of a chi-square test were already planned, so the finding of percentage rates which were not normally distributed did not require further action.
Chi-square analysis does not require normal distribution of variables or homogeneity of variances (Morgan et al., 2013). Chi-square analysis was performed using raw course completion numbers due to the issues with analyzing percentages, which was discussed in chapter 3.

Figure 4.3. Online and face-to-face completion rates of universities grouped by student classification. The scatterplot was produced using the percentages of students who completed courses and shows the higher completion rates of graduate students.

Inferential Statistics

Inferential statistics allow researchers to make inferences about a population based on sample data. Inferential statistics can be divided into difference and associational. This study
used difference inferential statistics, which is intended to compare differences and draw conclusions about statistical significances of the differences (Morgan et al., 2013).

**Contingency Table**

Research question 1: What is the difference among select four-year universities’ course completion rates of online distance learning courses across the regions of the United States?

Chi-square analysis was performed to determine if there was a statistically significant difference between each university and the online course completion. Assumptions were checked

Table 4.4

*Online Completers (n = 237,499) and Non-Completers (n = 34,210) by University for Fall 2013 and Spring 2014*

<table>
<thead>
<tr>
<th>University</th>
<th>Online Completers</th>
<th>Online Non-Completers</th>
<th>$X^2(1)$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>%</td>
<td>$n$</td>
<td>%</td>
</tr>
<tr>
<td>4A</td>
<td>6,357</td>
<td>2.7</td>
<td>908</td>
<td>2.7</td>
</tr>
<tr>
<td>5A</td>
<td>9,446</td>
<td>4.0</td>
<td>1,585</td>
<td>4.6</td>
</tr>
<tr>
<td>5B</td>
<td>91,101</td>
<td>38.4</td>
<td>12,934</td>
<td>37.8</td>
</tr>
<tr>
<td>6A</td>
<td>26,979</td>
<td>11.4</td>
<td>3,489</td>
<td>10.2</td>
</tr>
<tr>
<td>6B</td>
<td>28,455</td>
<td>12.0</td>
<td>5,546</td>
<td>16.2</td>
</tr>
<tr>
<td>6C</td>
<td>21,525</td>
<td>9.1</td>
<td>1,340</td>
<td>3.9</td>
</tr>
<tr>
<td>6D</td>
<td>5,996</td>
<td>2.5</td>
<td>428</td>
<td>1.3</td>
</tr>
<tr>
<td>7A</td>
<td>18,242</td>
<td>7.7</td>
<td>2,319</td>
<td>6.8</td>
</tr>
<tr>
<td>8A</td>
<td>14,169</td>
<td>6.0</td>
<td>2,066</td>
<td>6.0</td>
</tr>
<tr>
<td>8B</td>
<td>6,969</td>
<td>2.9</td>
<td>545</td>
<td>1.6</td>
</tr>
<tr>
<td>8C</td>
<td>8,260</td>
<td>3.5</td>
<td>3,050</td>
<td>8.9</td>
</tr>
<tr>
<td>Total</td>
<td>237,499</td>
<td>100</td>
<td>34,210</td>
<td>100</td>
</tr>
</tbody>
</table>
and met. Table 4.4 shows the university and chi-square results. The $X^2(1)$ and $p$ for each university indicate the degree of difference between the expected and actual course completion. The percentage shown for each university is another indicator of the difference between expected and actual course completion. A higher percentage of online completers than non-completers was a positive indication for that university.

Overall chi-square results indicated there was a statistically significant difference at the .05 level between university and online course completion, $X^2(10, n = 271,709) = 3,994.93, p < .001$, Cramér’s $V = .121$. Effect size showed the association strength to be small or weak (Cohen, 1988; Rea & Parker, 2014).

Research question 2: What is the difference in online distance learning course completion and face-to-face course completion among select four-year universities across the regions of the United States?

Chi-square analysis was performed to determine if there was a statistically significant difference between each university’s online and face-to-face course completion. Assumptions were checked and met. Table 4.5 shows the university and chi-square results. The $X^2(1)$ and $p$ for each university indicates the difference between the expected and actual course completion. The percentage shown for each university is the proportion of course completers to course enrollers. One university (6C) reported a higher percentage of online completers than face-to-face completers.

The chi-square analysis, shown in Table 4.5, indicated that overall there were more students who completed (2,288,007) face-to-face courses than expected (2,278,135) and fewer students who completed (237,499) the online courses than expected (247,371).
Overall, there was a statistically significant difference at the .05 level between face-to-face and online course completion, $\chi^2(1, n = 2,773,980) = 4,876.05, p < .001, \Phi = .042$. The effect size showed the difference to be very small or negligible (Cohen, 1988; Rea & Parker, 2014).

Table 4.5

*Online (n = 237,499) and Face-to-Face (n = 2,288,007) Course Completers by University for Fall 2013 and Spring 2014*

<table>
<thead>
<tr>
<th>University</th>
<th>Online Completers</th>
<th>Face-to-Face Completers</th>
<th>$\chi^2(1)$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4A</td>
<td>6,357 87.5</td>
<td>176,850 92.0</td>
<td>15.13</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>5A</td>
<td>9,446 85.6</td>
<td>103,993 88.5</td>
<td>9.08</td>
<td>.003</td>
</tr>
<tr>
<td>5B</td>
<td>91,101 87.6</td>
<td>270,420 87.9</td>
<td>.85</td>
<td>.357</td>
</tr>
<tr>
<td>6A</td>
<td>26,979 88.5</td>
<td>24,6245 91.3</td>
<td>23.41</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>6B</td>
<td>28,455 83.7</td>
<td>209,347 85.0</td>
<td>6.21</td>
<td>.013</td>
</tr>
<tr>
<td>6C</td>
<td>21,525 94.1</td>
<td>391,146 93.1</td>
<td>2.35</td>
<td>.126</td>
</tr>
<tr>
<td>6D</td>
<td>5,996 93.3</td>
<td>359,262 93.4</td>
<td>.01</td>
<td>.920</td>
</tr>
<tr>
<td>7A</td>
<td>18,242 88.7</td>
<td>155,381 90.6</td>
<td>7.33</td>
<td>.007</td>
</tr>
<tr>
<td>8A</td>
<td>14,169 87.3</td>
<td>46,542 93.6</td>
<td>52.72</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>8B</td>
<td>6,969 92.7</td>
<td>271,947 97.7</td>
<td>18.49</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>8C</td>
<td>8,260 73.0</td>
<td>56,874 87.8</td>
<td>244.05</td>
<td>&lt; .001</td>
</tr>
</tbody>
</table>
Research question 3: What is the difference in online distance learning course completion between undergraduate and graduate students at select four-year universities across the regions of the United States?

Chi-square analysis was performed to determine if there was a statistically significant difference between each university’s graduate and undergraduate course completion. Assumptions were checked and met. Table 4.6 shows the university and chi-square results. The $\chi^2(1)$ and $p$ for each university indicates the difference between the expected and actual course completion.

Table 4.6

*Undergraduate (n = 201,963) and Graduate (n = 35,536) Online Course Completers by University for Fall 2013 and Spring 2014*

<table>
<thead>
<tr>
<th>University</th>
<th>Undergraduate Completers</th>
<th>Graduate Completers</th>
<th>$\chi^2(1)$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>%</td>
<td>$n$</td>
<td>%</td>
</tr>
<tr>
<td>4A</td>
<td>4,699</td>
<td>84.7</td>
<td>1,658</td>
<td>96.6</td>
</tr>
<tr>
<td>5A</td>
<td>7,615</td>
<td>84.6</td>
<td>1,831</td>
<td>90.1</td>
</tr>
<tr>
<td>5B</td>
<td>81,733</td>
<td>87.1</td>
<td>9,368</td>
<td>92.2</td>
</tr>
<tr>
<td>6A</td>
<td>25,088</td>
<td>88.0</td>
<td>1,891</td>
<td>95.8</td>
</tr>
<tr>
<td>6B</td>
<td>21,379</td>
<td>81.6</td>
<td>7,076</td>
<td>90.7</td>
</tr>
<tr>
<td>6C</td>
<td>16,178</td>
<td>94.4</td>
<td>5,347</td>
<td>93.5</td>
</tr>
<tr>
<td>6D</td>
<td>4,893</td>
<td>92.3</td>
<td>1,103</td>
<td>98.5</td>
</tr>
<tr>
<td>7A</td>
<td>15,573</td>
<td>88.3</td>
<td>2,669</td>
<td>91.2</td>
</tr>
<tr>
<td>8A</td>
<td>12,581</td>
<td>86.8</td>
<td>1,588</td>
<td>91.4</td>
</tr>
<tr>
<td>8B</td>
<td>4,760</td>
<td>90.8</td>
<td>2,209</td>
<td>97.2</td>
</tr>
<tr>
<td>8C</td>
<td>7,464</td>
<td>71.9</td>
<td>796</td>
<td>86.2</td>
</tr>
</tbody>
</table>
completion. The percentage shown for each university is the proportion of course completers to course enrolers for that student classification. One university (6C) reported a higher percentage of undergraduate online completers than graduate online completers.

The chi-square analysis indicated that there were more graduate students (35,536) who completed online courses than expected (33,546) and fewer undergraduate students (201,963) who completed the online courses than expected (203,953).

Overall, there was a statistically significant difference at the .05 level between undergraduate and graduate online course completion, $X^2(1, n = 271,709) = 1,091.87, p < .001$, $\Phi = .063$. The effect size showed the difference to be very small or negligible (Cohen, 1988; Rea & Parker, 2014). University-level differences are shown in table 4.6.

**Multilevel Analysis**

Using generalized linear mixed model (GLMM) analysis to consider research question 2, the $z$-test ($z = 2.235, p = .025$) indicated the intercept variance varies between universities with statistical significance, which supports developing the multilevel model. The ICC is .083, or $0.298 / (0.298 + 3.29)$, which suggests about 8.3% of the variability in course completion lies between universities. Using the null model shown in Table 4.7, the estimated logged event rate for an undergraduate face-to-face student completing a course is 2.353 (where the random effect is at 0 or is the average variability among universities.)

After adding fixed and random effects for the final model ($\beta = 2.33, SE = .15, p < .001$), shown in Tables 4.7 and 4.8, the estimated logged event rate for an undergraduate face-to-face student completing a course is 2.333. Both level 1 predictors, online and graduate, are significantly related at the .05 level to a student’s probability of completing a course. The model indicated a negative effect for the online student and a positive effect for the graduate student.
The ratio of difference in expected event rates between online and face-to-face students of .67, or 6.909/10.312, suggest that for online students, the estimated count decreased by about 33% compared with the estimated count of face-to-face students. Graduate students can be expected to have an event rate 2.08 times greater for course completion compared to undergraduate students, holding all other variables constant. The $z$-test ($z = 2.235, p = .025$) remained unchanged, indicating the intercept variance varies between universities with statistical significance after adding within-university predictors.

Table 4.7.

_Online and Face-to-Face Students: GLMM logistic parameter estimates (Est.), standard errors (SE) and p Values_

<table>
<thead>
<tr>
<th>Effect</th>
<th>Null with Random Intercepts Model</th>
<th>Fixed Effects Model with Random Intercepts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>SE</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>2.353</td>
<td>.157</td>
</tr>
<tr>
<td>Online</td>
<td>- .400</td>
<td>.119</td>
</tr>
<tr>
<td>Graduate</td>
<td>.733</td>
<td>.144</td>
</tr>
<tr>
<td>Random Effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between-univ</td>
<td>.298</td>
<td>.133</td>
</tr>
<tr>
<td>Graduate</td>
<td>.251</td>
<td>.113</td>
</tr>
<tr>
<td>Online</td>
<td>.170</td>
<td>.077</td>
</tr>
<tr>
<td>Information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criteria (BIC)</td>
<td>20,511.629</td>
<td></td>
</tr>
</tbody>
</table>
The odds ratio for online students completing a course was 1.49, or $2.71829^{-0.40}$, and 2.08, or $2.71829^{0.733}$, for graduates. Converting to Cohen’s $d$ results in .22 and .40 respectively, which can both be considered small (Cohen, 1988); graduate effect size can be considered educationally significant (Spybrook, 2008).

Table 4.8

*Fixed Effects with Random Intercepts Model*

<table>
<thead>
<tr>
<th>Effect</th>
<th>Est.</th>
<th>SE</th>
<th>$t/Z$</th>
<th>$p$ value</th>
<th>Exp</th>
<th>95% Confidence Interval for Exponentiated Coeff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online</td>
<td>-.400</td>
<td>.119</td>
<td>-3.362</td>
<td>.001</td>
<td>.670</td>
<td>.529 – .849</td>
</tr>
<tr>
<td>Graduate</td>
<td>.733</td>
<td>.144</td>
<td>5.080</td>
<td>&lt;.001</td>
<td>2.082</td>
<td>1.562 – 2.774</td>
</tr>
<tr>
<td>Random Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between-univ</td>
<td>.277</td>
<td>.124</td>
<td>2.235</td>
<td>.025</td>
<td>.115</td>
<td>.115 – .665</td>
</tr>
<tr>
<td>Graduate</td>
<td>.251</td>
<td>.113</td>
<td>2.228</td>
<td>.026</td>
<td>.104</td>
<td>.104 – .606</td>
</tr>
<tr>
<td>Online</td>
<td>.170</td>
<td>.077</td>
<td>2.219</td>
<td>.027</td>
<td>.070</td>
<td>.070 – .412</td>
</tr>
</tbody>
</table>

In considering research questions 1 and 3, the $z$-test ($z = 2.227$, $p = .026$) indicated the intercept variance varies between universities with statistical significance, which supports developing the multilevel model. The ICC is .071, or $.250 / (.250 + 3.29)$, which suggests about 7.1% of the variability in online course completion lies between universities. The estimated
logged event rate for an undergraduate online student completing a course is 2.028 (where the random effect is at 0 or the average variability among universities.)

After adding fixed and random effects for the final model ($\beta = 1.93$, SE = .14, $p < .001$), shown in Tables 4.9 and 4.10, the estimated logged event rate for an undergraduate online student completing a course is 1.930 (about 93%). The level 1 predictor, graduate, is significantly related at the .05 level to a student’s probability of completing a course. The model indicated a positive effect for the graduate student.

Table 4.9

*Online Students: GLMM logistic parameter estimates (Est.), standard errors (SE) and p Values*

<table>
<thead>
<tr>
<th>Effect</th>
<th>Null Model with Random Intercepts</th>
<th>Fixed Effects Model with Random Intercepts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est</td>
<td>SE</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>2.028</td>
<td>.144</td>
</tr>
<tr>
<td>Graduate</td>
<td>.813</td>
<td>.160</td>
</tr>
<tr>
<td>Random Effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between-univ</td>
<td>.250</td>
<td>.112</td>
</tr>
<tr>
<td>Graduate</td>
<td>.296</td>
<td>.141</td>
</tr>
<tr>
<td>Information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criteria (BIC)</td>
<td>1,237.073</td>
<td>91.200</td>
</tr>
</tbody>
</table>

The ratio of difference in expected event rates between online graduate and online undergraduate students of 3.113, or 28.699/9.219, suggests that for graduate students, the
estimated count increased by about 200% compared with the estimated count of undergraduate students. The z-test \((z = 2.227, p = .026)\) remains unchanged, indicating the intercept variance varies between universities with statistical significance after adding within-university predictors. The odds ratio for online graduate students completing a course is 2.25, or \(2.71829^{.813}\). Converting to Cohen’s \(d (.44)\) indicates a small effect size (Cohen, 1988), which can be considered educationally significant (Spybrook, 2008).

Table 4.10

*Fixed Effects with Random Intercepts Model (Binomial Probability Distribution, Logit Link Function)*

<table>
<thead>
<tr>
<th>Effect</th>
<th>Est.</th>
<th>SE</th>
<th>(t/Z)</th>
<th>(p) value</th>
<th>Exp</th>
<th>95% Confidence Interval for Exponentiated Coeff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.930</td>
<td>.144</td>
<td>13.388</td>
<td>&lt;.001</td>
<td>6.892</td>
<td>5.152 – 9.219</td>
</tr>
<tr>
<td>Graduate</td>
<td>.813</td>
<td>.160</td>
<td>5.090</td>
<td>&lt;.001</td>
<td>2.255</td>
<td>1.634 – 3.113</td>
</tr>
<tr>
<td>Random Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between-univ</td>
<td>.250</td>
<td>.112</td>
<td>2.227</td>
<td>.026</td>
<td>.104</td>
<td>.104 – .604</td>
</tr>
<tr>
<td>Graduate</td>
<td>.296</td>
<td>.141</td>
<td>2.109</td>
<td>.035</td>
<td>.117</td>
<td>.117 – .751</td>
</tr>
</tbody>
</table>

**Summary**

Eleven universities provided course completion data for two semesters, Fall 2013 and Spring 2014, by course mode and student classification. The total enrollment, \(n = 2,525,506\)
was not the actual number of students, but the number of courses students enrolled in. A purposive sample of large, high or very high research universities were invited to participate in the study.

Descriptive statistics for percentages and enrollment numbers were both presented, but analysis was conducted only using the enrollment and course completion numbers. Analysis of percentages can be problematic unless the denominator is the same for all cases. Percentages often do not add up correctly when calculating at sub-levels, nor do they necessarily represent the population size.

Analysis was first conducted using chi-square test for independence or association to answer the three research questions. All questions were considered statistically significant with an effect size ranging from negligible to small. In order to account for group effects, further analysis was pursued using a type of multilevel modeling called generalized linear mixed modelling (GLMM). The results were also statistically significant for all three research questions, but this time with an effect size of a higher magnitude, ranging from small to small but educationally significant. The next chapter summarizes and discusses these results.
CHAPTER 5
SUMMARY AND DISCUSSION

The purpose of this study was to compare retention or course completion in online learning and face-to-face learning at select four-year public universities. The study was conducted because online higher education has a reputation for having a lower course completion rate than traditional face-to-face education, but the previous research supporting this belief is limited to individual courses, small institutions, anecdotal reports, or is outdated.

The goal of this study was not to prove retention is or is not lower in online courses than in face-to-face courses. The intent was to improve upon statements in existing research that describe a difference in retention or course completion between online learning and face-to-face learning among institutions. Leading students and instructors to believe that online students drop or fail their courses more often than face-to-face students can contribute to a higher dropout or failure rate by acting as a self-fulfilling prophecy (Rosenthal, 1976). This possibility leads to a concern that online students may be predisposed to dropping out if they think everyone else in an online course is dropping out. This study indicated that not everyone in online classes drops out or fails—and the completion rate is not much different than for face-to-face classes.

Retention is widely considered to be an indicator of the quality of instruction, both in online and face-to-face education. This has not always been the case. A high dropout rate used to be an indicator of a challenging course or institution. Now a high dropout rate is often viewed as a lost customer and an issue which needs fixed. Much research in online education has been devoted to identifying or validating the factors leading to dropout or failure or on changing the courses using the results of these studies.

The research questions this study was designed to answer were:
1. What is the difference among select four-year universities’ course completion rates of online distance learning courses across the regions of the United States?

2. What is the difference in online distance learning course completion and face-to-face course completion among select four-year universities across the regions of the United States?

3. What is the difference in online distance learning course completion between undergraduate and graduate students at select four-year universities across the regions of the United States?

Eleven large, highly or very highly productive public research universities provided course completion information for the Fall 2013 and Spring 2014 semesters. One-third of the total number universities solicited chose to participate in the study. Course completion rate was used in place of retention rate for this study. Retention rates, as often used, do not distinguish between online and face-to-face students, but instead considers the progress of entire cohorts of degree-seeking students. Course completion is a necessary component of retention that allows for comparison of online and face-to-face courses. Students who successfully complete courses become students who are retained through to graduation. Course completion information consisted of the number of students who enrolled in courses and the number of students who completed courses successfully. Students enrolled in multiple courses were counted for each course. Course enrollments totaled 2,502,271 for face-to-face students for two semesters, with 2,288,007 students evaluated as successfully completing courses. Online student course enrollments totaled 271,709 students for two semesters with 237,499 students considered to be successful completers.

The results indicated course completion rates as a percentages were higher for face-to-face courses at all but one of the eleven participating universities. However, the findings indicated the difference between online and face-to-face course completion should be considered
of low practical significance ($\beta = -.40, p = .001, d = .22$). Analysis of the university online course completion rates indicated the differences between universities could also be considered of not much practical significance ($X^2 = 3,994.93, p < .001$, Cramér’s $V = .121$). The results also indicated undergraduate online student course completion was somewhat lower than graduate online course completion, showing an effect size often considered to have educational significance ($\beta = .81, p < .001, d = .44$). In education, even an effect size as small as .20 or .30 can be considered important enough to pay attention to and consider making a change, but other factors need evaluated as well.

The results of this study were different from studies found in the literature in that most did not report effect size, which is an indicator of practical significance. The study demonstrated that online course completion or retention may not have the magnitude of the problem it is often credited with. A pessimist can view these results as typical for online instruction. An optimist can view these results as better than they had hoped for and that online completion rates could have been much lower. A statistician and a researcher can consider the importance of effect size, recognize that the practical significance is considered small for two of the three research questions and educationally significant for one research question, and reflect upon possible follow-on studies. Interpretation of results are discussed in more detail in the next sections, which are organized by research question.

Research Question 1

The first question asked about the difference among select four-year universities’ course completion rates of online distance learning courses across the regions of the United States; while there was a difference among university’s online course completion rates, the practical significance of the difference was small ($X^2(10, n = 271,709) = 3,994.93, p < .001$, Cramér’s $V = $
This question was answered by looking at enrollment instances of 271,709 online students across eleven (11) participating universities. The main purpose of the research question was to identify the universities with the highest (94%) and lowest (73%) online completion rates to select students from in order to conduct interviews in a follow-on phase. Lessons should be sought out and learned from these two universities; the numbers alone do not tell the story.

Looking among the eleven universities, the analysis showed online course completion rates were higher than expected at six universities and lower at three. For the lower three universities, implications of these rates are not clear. The results could mean students from those universities dropped or failed courses at a higher rate than the other eight universities in the study. It could also simply mean the three universities calculated the completion rates differently than the other universities.

Research Question 2

The second question focused on the difference in online distance learning course completion and face-to-face course completion among select four-year universities across the regions of the United States. Results indicated there was a likelihood for online and face-to-face modes to differ in course completion; however, the practical significance or effect size was small or negligible ($\beta = -.40$, $SE = .12$, $p = .001$, $d = .22$). This question was answered by looking at enrollment instances of 2,773,980 students in both online and face-to-face courses among eleven (11) participating universities. Online enrollments ($n = 271,709$) were about 11% of the number of face-to-face enrollments ($n = 2,502,271$). One university reported a higher completion rate for online students than face-to-face students; by contrast, all other universities reported a lower percentage for online completion. Universities that tended to have higher online course completion rates than other universities also tended to have higher face-to-face course
completion rates than other universities. The same was true for those with lower online and face-to-face retention rates.

Research Question 3

The final question asked about the difference in online distance learning course completion between undergraduate and graduate students at select four-year universities across the regions of the United States. Results indicated graduate students were more likely than undergraduate students to complete an online course. Effect size was either negligible or small but educationally significant, depending on the type of analysis considered ($\beta = .81, SE = .16, p < .001, d = .44$). This question was addressed by examining the enrollment instances of 271,709 students in online courses across the participating universities. In the undergraduate/graduate online completion comparison between universities, all but one university had undergraduate completion numbers that were lower than expected, while the graduate online completion numbers were higher than expected. This indicates that the bulk of the non-completers in the entire study were contained in the undergraduate online courses. This could be an indication of the impact of student age on the level of responsibility needed to succeed in an online course.

Relationship of the Current Study to Previous Research

This study expands upon and update existing research and may reflect changes in students, technology, and its possible influences upon online students. Since the most commonly cited references for lower online retention or course completion are from Carr (2000), Frankola (2001), and Waschull (2001) and were published almost fifteen years ago, technology has improved in the form of Web 2.0 collaborative tools and expanded broadband access. These changes should not be ignored even though Xu & Jaggars (2011) suggested that technological improvements from 2004 to 2008 did not have an impact on online student success in their study,
though that may be partly due to lack of use of expertise on the part of instructors or students due to their recency, novelty, or specific implementations.

This study builds on and broadens two of the three studies found in the literature that reported effect size, or practical significance. This study’s finding with the most practical significance indicated that undergraduate students failed to complete their course more often than graduate students, which most directly expounds upon and confirms an earlier study (Xu & Jaggers, 2011). That study found online community college students in freshman level courses dropped out more than their face-to-face counterparts with an educationally significant effect size. Results from both studies can be an indication that an online course is not the best fit for a beginning college student.

The comparison of online course completion with face-to-face course completion rates is similar to a study by Patterson and McFadden (2009) in that a difference was found and effect size was reported, but this current study was larger in scope and also included undergraduate students. In the reviewed literature, most of these studies were of causal-comparative design and used chi-square analysis, which should have made for a sound comparison to this study. Results of the comparisons often did not address practical significance or effect size but did show statistical significance. Failing to report effect size can lead to a different perception of the effectiveness of online instruction and does not tell the story as accurately as it should.

Relationship of the Current Study to the Theoretical Framework

As this study indicates, undergraduate online students had a higher propensity to not complete an online course than graduate online students. This propensity can be viewed through the lens of the theoretical framework, Rovai’s (2003) composite persistent model, as suggesting student age, intellectual development, and academic preparation may need further consideration.
These are all likely differences between undergraduate and graduate students if viewed from Rovai’s (2003) model. However, students possess these characteristics prior to admission, which may indicate ungraduated students are going to be more predisposed towards not completing an online course than graduate students.

Undergraduate and graduate students could differ on other factors in the model that need to be considered after admission, such as the internal factor of student needs. The student needs of self-esteem and interpersonal relationships likely are different between a student who is an undergraduate and one who is a graduate. The undergraduates and graduates may be impacted differently by whether an online course includes a synchronous component or an asynchronous component. Questions remain about external factors which may differ between undergraduate and graduate students, to include family responsibilities and hours of employment. Administrators, instructors, and instructional designers may wish to consider differences in student classification and their likely tendencies when involved with undergraduate online students and courses.

Limitations of the Study

There were two original limitations of the study which are presented in chapter one. First, it could be argued that it was inappropriate to compare face-to-face course completion or retention with online course completion since the courses commonly attract different student audiences. Secondly, secondary data was drawn from different sources and there is a chance the data providers interpreted the request for data differently than intended and a university was represented incorrectly.

Some consider hypothesis testing a limitation because it is often not taught or carried out the way the original authors intended (Kline, 2013). While researching the issues with hypothesis
testing, Bayesian statistics kept showing up as a possible alternative to hypothesis testing. This is a topic which could be explored in future research, but it appears that using different statistics methods would make it more difficult to compare to other research studies and for other research studies to compare to this one.

Other limitations are related to the data received from the universities; namely, the request for data could have been interpreted differently than intended and not actually have been adequate for comparing to other universities. For example, one university reported counting all Ds as failures, but the university policy for counting a D as a failure depended on the student and the program itself. Another university claimed to remove duplicate entries caused by students taking more than one course. This should have led to a lower enrollment than most other universities reported, but it appeared not to. There are most likely other differences in the data which remain limitations. Having a limitation such as these does not remove the value obtained by conducting the research.

**Discussion of the Results**

The findings support Clark’s (1983) position that the media does not matter due to the small practical significance of the difference between online and face-to-face completion rates. There are researchers who believe online and face-to-face courses should not be compared to each other at all, and those who caution that the differences between online and face-to-face instruction are about more than just the media. Comparisons between the two course modes will most likely continue, as they should. This study can provide researchers an updated, comprehensive, comparative analysis between online and face-to-face course completion or retention rates.
Using a metaphorical situation, if comparing between two face-to-face classrooms with different instructors using the same media tools, one instructor would likely be rated higher than the other. This is not a reason to remove the instructor, nor to consider the lower rated instructor inferior. There are too many other factors involved, such as students in the classroom that change every semester and the course content. There may be quite a few students who prefer the lower rated instructor over the higher rated instructor, but the students of the higher rated instructor may be more vocal about their praise or were simply pushed to complete the course evaluation by a particular instructor. If the difference continued over a period of years, the opinion of the instructor’s qualifications would not change since the difference is of a small magnitude—or would the instructor in fact be relieved of duties if the statistic used to rate him or her was always a few percentage points lower than another instructor? It makes more sense to consider that the other instructor was slightly better on that statistic and not penalize the instructor with the slightly lower rating; that instructor is still good enough to be entrusted with educating students. Not everyone can be the top-rated instructor. If there is really only room for one top-rated instructor, the standards need to change because students need more than one instructor. The statistics used to evaluate the instructors do serve a purpose and should not be discontinued. They most likely serve as an incentive or encouragement to both instructors to continue improving their skills.

Unexpected Findings

The apparent low quality of research comparing online and face-to-face course completion rates was a surprise. Most of the studies found within the last ten years were dissertations and did not report effect size, or practical significance. Effect size reporting has been considered a basic requirement for publishing since 2001. One could speculate the authors
intentionally left out the effect size because of a small practical significance. This would not be a logical occurrence with a dissertation, though; a low effect size is widely considered a drawback to a study which can limits an author’s ability to get published. This finding is similar to that of Bernard et al. (2004) who conducted a meta-analysis of 232 research studies comparing distance education with face-to-face education and criticized the studies for being of “poor methodological quality and severely lacking in critical information about research practices” (p. 175).

The analysis also demonstrated the importance of the selection of statistical procedures along with the reporting of effect size. The data collected were not restricted to one method of analysis; other methods could be performed which might lead to different results. Use of ratios and groups warranted careful consideration of analysis methods. A researcher should use caution in selecting the statistical method when analyzing proportions in particular. Non-researchers should use caution as well and look beyond the proportions during decision making. It is not enough to determine which group or treatment has a higher percentage in the results. Percentages alone do not take differences in group sizes into consideration and can be misleading. In order to account for percentages in this study, chi-square analysis and generalized linear multilevel modeling (GLMM) were both used. Educational researchers considering analysis involving percentages or including students in more than one group, such as multiple classrooms, districts, or schools, can account for variances caused by the different groups or populations by using GLMM or another type of multilevel analysis.

Finding one university with a higher online than face-to-face course completion rate was not an anticipated result. The highest anticipated result was a finding of no statistically significant difference, with face-to-face still having a higher completion rate than online. One
has to wonder if the findings of no statistically significance difference have impacted what is published in the peer-reviewed journals, but this finding should be what online and distance learning journals would want to publish.

**Suggestions for Future Research**

Future researchers should consider ways to improve online instruction, of course, but it should be pursued without basing the need upon a shaky claim, such as that online retention is lower than face-to-face retention. Online instruction should be improved for the good of online students, instructors, and society.

This research study could be considered part of a larger setting of exploratory research. Stebbins (2001) defined exploratory research in the social sciences as “a broad-ranging, purposive, systematic, prearranged undertaking designed to maximize the discovery of generalizations leading to description and understanding of an area of social or psychological life” (p. 3). It may be fitting to continue conducting exploratory research to compare online and face-to-face course completion. Stebbins’ (2001) discussion of exploratory research is appropriate to also describe the comparison between online and face-to-face course completion as found in the literature review:

In general, exploration is the preferred methodological approach under at least three conditions: when a group, process, activity, or situation has received little or no systematic empirical scrutiny, has been largely examined using prediction and control rather than flexibility and open-mindedness, or has grown to maturity . . . but has changed so much along the way that it begs to be explored anew. (p. 9)

Future research is planned that would make this a mixed methods study. One intent of the mixed methods study is to conduct interviews of students from the two universities with the highest and lowest online course completion rates, known as extreme case sampling. Other results from this study could be explored as well. The educational significance of the difference between
undergraduate and graduate online course completion could be researched further by looking at factors from Rovai’s (2003) composite persistence model, as discussed earlier in this chapter.

Similar studies that are national in nature such as this one should be pursued by other researchers. This study looked at a large population by examining only large, high or very high research, public universities. A nationwide study of community colleges could likely result in an even larger number of enrollments to analyze. Other university populations could be examined, such as large universities that are not research-intensive. Additional research could be pursued with a large nationwide undergraduate population in multiple college settings which includes more variables, such as year in college and major or program of study and some student characteristics, such as age or year in college. Individual characteristics can indicate, for example, that student age or year in college is the variable influencing course completion rates in online instruction more than student classification (undergraduate or graduate) alone. Further research may help narrow down the specifics of the undergraduate student population so efforts can be focused on either changing the specific courses which are offered to undergraduate students or changing the content of the courses offered.

Follow-on research can be pursued with the existing data in this study. University characteristics can be obtained from a publicly available source, to include IPEDS and Carnegie data, and variables added to the model, such as acceptance rate, average SAT/ACT score, or locale (city/town/suburb/rural). Adding more university-level characteristics would make the use of multilevel analysis more appropriate and provide additional, important context and improve the utility of the research outcomes.

Conclusion
As stated, online completion was lower than face-to-face completion in this study, with a small practical significance. Online completion was often lower in previous studies, but this should be expected and is not a reason to eliminate or discredit online instruction. Distance education has been in widespread use for over 120 years; however, technology has helped it grow and improve, especially in the past ten years. There was no expectation of finding a higher retention rate in online instruction. Using that as a goal for online retention is not realistic and can be considered as discouraging as a self-fulfilling prophecy can be to a student. Researchers and practitioners should continue to find ways to improve online instruction, just as they should continue to find ways to improve face-to-face instruction. Researchers of online instruction should have a legitimate reference to use as they continue studying ways to improve online education. Referring to a ten-year old newspaper article based on anecdotal information discredits their study. This study fills a current gap in the literature. Researchers may wish to use this study as a reference to demonstrate online retention is lower than face-to-face retention—even if the effect size is considered small—if that fits their purpose. Given the study involved over two million student course enrollments and eleven large universities, it may be a better reference than the studies found in the literature review contained here.

Other student characteristics such as gender and ethnicity have been explored more in research than the overall comparison of online to face-to-face retention. Those factors or variables should be studied, but they should not be researched based on the claim that online retention is lower than face-to-face retention as that can add bias to the study. Just as we do not intend to change student gender or ethnicity to fit online instruction, we should not change the integrity of our research to discredit online instruction. We also should not expect students to change their preference or need for online or face-to-face instruction by making articles available
for them to read that use outdated or incomplete information on the quality of either mode of course instruction. Students should expect high quality research to read and learn from, just as they should expect high quality instruction from the course mode they participate in, whether it is online or face-to-face.
APPENDIX A

SAMPLE E-MAIL REQUEST
Sample of the e-mail request for data sent to universities:

Dr. Doe, I am looking for Your University enrollment and course completion data for Fall 2013 and Spring 2014, broken out by face-to-face and online course and undergraduate and graduate level. (Spreadsheet attached).

I am a doctoral candidate at the University of North Texas, majoring in Learning Technologies. I am conducting research for my dissertation titled “Retention: Course Completion Rates in Online Distance Learning”. The purpose of this study will be to compare retention or course completion in online learning and face-to-face learning at select four-year public universities.

Three to four universities from eight IPEDS regions will be included in the study; university data will be anonymized and they will be identified by region only.

Please let me know if you need any other details to consider this request.

Thank you in advance for your time!

--Alana S. Phillips
APPENDIX B

SPREADSHEET SHELL SENT TO UNIVERSITIES
Spreadsheets shell sent to universities

<table>
<thead>
<tr>
<th>University Name</th>
<th>Semester</th>
<th>Course Mode</th>
<th>Student Classification</th>
<th># Students Enrolled at Census Date</th>
<th># Students Successfully Completing Course</th>
<th>Course Completion Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fall 2013</td>
<td>Online Course</td>
<td>Undergraduate</td>
<td>#DIV/0!</td>
<td>#DIV/0!</td>
<td>#DIV/0!</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Graduate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Face-to-face Course</td>
<td>Undergraduate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Face-to-face Course</td>
<td>Graduate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spring 2014</td>
<td>Online Course</td>
<td>Undergraduate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Graduate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Face-to-face Course</td>
<td>Undergraduate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Face-to-face Course</td>
<td>Graduate</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Definitions:

*Course completion* can be defined differently by institutions, but generally is receiving a passing grade in a course. Students who complete a course unsuccessfully are counted in the failure category. Some institutions or programs count a D as a failure.

*Face-to-face course* can be defined differently by institutions, but generally includes courses where there is no technological content or up to 29% of the content may be delivered online. These can be considered either traditional or web-facilitated courses (Allen & Seaman, 2013).

*Online course* can be defined differently by institutions, but generally is a course where the instructor normally delivers over 80% (Allen & Seaman, 2013) of the “content, instruction, and materials over the Internet and the student attends class within this online classroom” (Robinson et al., 2015, p. 58). Moore and Kearsley (2011) emphasized that an online course meets at a different time or location or both, so the technology is not what makes it an online course, it is how the instructor meets with the students.
APPENDIX C

T-TEST OF CONVERTED AND ACTUAL PERCENTAGES
### T-Test of Converted and Actual Percentage

#### Group Statistics

<table>
<thead>
<tr>
<th>Course Mode</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARCSINE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face-to-face</td>
<td>44</td>
<td>1.307</td>
<td>0.086</td>
<td>0.013</td>
</tr>
<tr>
<td>Online</td>
<td>44</td>
<td>1.257</td>
<td>0.099</td>
<td>0.015</td>
</tr>
<tr>
<td>logit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face-to-face</td>
<td>44</td>
<td>1.190</td>
<td>0.343</td>
<td>0.052</td>
</tr>
<tr>
<td>Online</td>
<td>44</td>
<td>1.021</td>
<td>0.320</td>
<td>0.048</td>
</tr>
<tr>
<td>Rate Calculated</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face-to-face</td>
<td>44</td>
<td>92.576</td>
<td>4.251</td>
<td>0.641</td>
</tr>
<tr>
<td>Online</td>
<td>44</td>
<td>89.713</td>
<td>5.949</td>
<td>0.897</td>
</tr>
</tbody>
</table>
### Independent Samples Test

<table>
<thead>
<tr>
<th></th>
<th>Levene’s Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
<td>t</td>
</tr>
<tr>
<td>ARCSINE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>0.11</td>
<td>0.741</td>
<td>2.542</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>2.542</td>
<td>84.54</td>
<td>0.013</td>
</tr>
<tr>
<td>logit</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>0.627</td>
<td>0.431</td>
<td>2.386</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>2.386</td>
<td>85.593</td>
<td>0.019</td>
</tr>
<tr>
<td>RateCalculated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal variances assumed</td>
<td>1.76</td>
<td>0.188</td>
<td>2.597</td>
</tr>
<tr>
<td>Equal variances not assumed</td>
<td>2.597</td>
<td>77.835</td>
<td>0.011</td>
</tr>
</tbody>
</table>
APPENDIX D

SHAPIRO-WILK TEST OF NORMALITY
### Shapiro-Wilk test of Normality

<table>
<thead>
<tr>
<th>Universities</th>
<th>Shapiro-Wilk test of normality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic</td>
</tr>
<tr>
<td>4A</td>
<td>.851</td>
</tr>
<tr>
<td>5A</td>
<td>.959</td>
</tr>
<tr>
<td>5B</td>
<td>.812</td>
</tr>
<tr>
<td>6A</td>
<td>.792</td>
</tr>
<tr>
<td>6B</td>
<td>.880</td>
</tr>
<tr>
<td>6C</td>
<td>.864</td>
</tr>
<tr>
<td>6D</td>
<td>.787</td>
</tr>
<tr>
<td>7A</td>
<td>.880</td>
</tr>
<tr>
<td>8A</td>
<td>.931</td>
</tr>
<tr>
<td>8B</td>
<td>.859</td>
</tr>
<tr>
<td>8C</td>
<td>.725</td>
</tr>
</tbody>
</table>
APPENDIX E

COURSE COMPLETION RATES FOR UNIVERSITIES BY REGION
Course completion rates for universities for each semester by region.

<table>
<thead>
<tr>
<th>Region</th>
<th>University</th>
<th>Semester</th>
<th>Course Mode</th>
<th>Student Classification</th>
<th># Students Enrolled</th>
<th># Students Completing Course</th>
<th>Completion Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>4A</td>
<td>Fall 2013</td>
<td>Online Course</td>
<td>Undergraduate</td>
<td>2250</td>
<td>1859</td>
<td>82.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Graduate</td>
<td>841</td>
<td>810</td>
<td>96.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Face-to-face Course</td>
<td>88245</td>
<td>80274</td>
<td>90.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Undergraduate</td>
<td>12433</td>
<td>11980</td>
<td>96.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Graduate</td>
<td>875</td>
<td>848</td>
<td>96.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spring 2014</td>
<td>Online Course</td>
<td>Undergraduate</td>
<td>3299</td>
<td>2840</td>
<td>86.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Graduate</td>
<td>875</td>
<td>848</td>
<td>96.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Face-to-face Course</td>
<td>79737</td>
<td>73027</td>
<td>91.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Undergraduate</td>
<td>11907</td>
<td>11569</td>
<td>97.16</td>
</tr>
<tr>
<td>5</td>
<td>5A</td>
<td>Fall 2013</td>
<td>Online Course</td>
<td>Undergraduate</td>
<td>4275</td>
<td>3671</td>
<td>85.87</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Graduate</td>
<td>1066</td>
<td>968</td>
<td>90.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Face-to-face Course</td>
<td>54792</td>
<td>48229</td>
<td>88.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Undergraduate</td>
<td>10775</td>
<td>10057</td>
<td>93.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spring 2014</td>
<td>Online Course</td>
<td>Undergraduate</td>
<td>4724</td>
<td>3944</td>
<td>83.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Graduate</td>
<td>966</td>
<td>863</td>
<td>89.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Face-to-face Course</td>
<td>43366</td>
<td>37622</td>
<td>86.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Undergraduate</td>
<td>8640</td>
<td>8085</td>
<td>93.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Graduate</td>
<td>8640</td>
<td>8085</td>
<td>93.58</td>
</tr>
<tr>
<td>5B</td>
<td>Fall 2013</td>
<td></td>
<td>Online Course</td>
<td>Undergraduate</td>
<td>45609</td>
<td>39774</td>
<td>87.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Graduate</td>
<td>5170</td>
<td>4753</td>
<td>91.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Face-to-face Course</td>
<td>146354</td>
<td>127404</td>
<td>87.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Undergraduate</td>
<td>13208</td>
<td>12720</td>
<td>96.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spring 2014</td>
<td>Online Course</td>
<td>Undergraduate</td>
<td>48270</td>
<td>41959</td>
<td>86.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Graduate</td>
<td>4986</td>
<td>4615</td>
<td>92.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Face-to-face Course</td>
<td>136004</td>
<td>118535</td>
<td>87.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Undergraduate</td>
<td>12155</td>
<td>11761</td>
<td>96.76</td>
</tr>
<tr>
<td>6</td>
<td>6A</td>
<td>Fall 2013</td>
<td>Online Course</td>
<td>Undergraduate</td>
<td>13638</td>
<td>12016</td>
<td>88.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Graduate</td>
<td>1004</td>
<td>964</td>
<td>96.02</td>
</tr>
<tr>
<td>Course Type</td>
<td>Format</td>
<td>Undergraduate</td>
<td>Graduate</td>
<td>Pass Rate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------</td>
<td>---------------</td>
<td>----------</td>
<td>-----------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face-to-face Course</td>
<td>Undergraduate</td>
<td>114574</td>
<td>103818</td>
<td>90.61</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring 2014</td>
<td>Graduate</td>
<td>25462</td>
<td>24411</td>
<td>95.87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online Course</td>
<td>Undergraduate</td>
<td>14856</td>
<td>13072</td>
<td>87.99</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>Graduate</td>
<td>970</td>
<td>927</td>
<td>95.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face-to-face Course</td>
<td>Undergraduate</td>
<td>104556</td>
<td>94062</td>
<td>89.96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6B Fall 2013</td>
<td>Graduate</td>
<td>25003</td>
<td>23954</td>
<td>95.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online Course</td>
<td>Undergraduate</td>
<td>13119</td>
<td>10768</td>
<td>82.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>Graduate</td>
<td>3931</td>
<td>3574</td>
<td>90.92</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face-to-face Course</td>
<td>Undergraduate</td>
<td>117774</td>
<td>99437</td>
<td>84.43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring 2014</td>
<td>Graduate</td>
<td>11294</td>
<td>10742</td>
<td>95.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online Course</td>
<td>Undergraduate</td>
<td>13082</td>
<td>10611</td>
<td>81.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>Graduate</td>
<td>3869</td>
<td>3502</td>
<td>90.51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face-to-face Course</td>
<td>Undergraduate</td>
<td>106561</td>
<td>89088</td>
<td>83.60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6C Fall 2013</td>
<td>Graduate</td>
<td>10614</td>
<td>10080</td>
<td>94.97</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online Course</td>
<td>Undergraduate</td>
<td>7410</td>
<td>6949</td>
<td>93.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>Graduate</td>
<td>2796</td>
<td>2603</td>
<td>93.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face-to-face Course</td>
<td>Undergraduate</td>
<td>195116</td>
<td>180148</td>
<td>92.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring 2014</td>
<td>Graduate</td>
<td>24583</td>
<td>23949</td>
<td>97.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online Course</td>
<td>Undergraduate</td>
<td>9734</td>
<td>9229</td>
<td>94.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>Graduate</td>
<td>2925</td>
<td>2744</td>
<td>93.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face-to-face Course</td>
<td>Undergraduate</td>
<td>179044</td>
<td>166413</td>
<td>92.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6D Fall 2013</td>
<td>Graduate</td>
<td>21240</td>
<td>20636</td>
<td>97.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online Course</td>
<td>Undergraduate</td>
<td>2295</td>
<td>2098</td>
<td>91.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>Graduate</td>
<td>494</td>
<td>488</td>
<td>98.79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face-to-face Course</td>
<td>Undergraduate</td>
<td>170552</td>
<td>158022</td>
<td>92.65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring 2014</td>
<td>Graduate</td>
<td>28653</td>
<td>28062</td>
<td>97.94</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online Course</td>
<td>Undergraduate</td>
<td>3009</td>
<td>2795</td>
<td>92.89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>Graduate</td>
<td>626</td>
<td>615</td>
<td>98.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face-to-face Course</td>
<td>Undergraduate</td>
<td>159620</td>
<td>147869</td>
<td>92.64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Fall 2013</td>
<td>Graduate</td>
<td>25687</td>
<td>25309</td>
<td>98.53</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online Course</td>
<td>Undergraduate</td>
<td>8442</td>
<td>7455</td>
<td>88.31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>Graduate</td>
<td>1527</td>
<td>1362</td>
<td>89.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Course Type</td>
<td>Year</td>
<td>Quarter</td>
<td>Course Type</td>
<td>Undergraduate</td>
<td>Graduate</td>
<td>Undergraduate</td>
<td>Graduate</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------</td>
<td>-------------</td>
<td>---------------------</td>
<td>---------------</td>
<td>----------</td>
<td>---------------</td>
<td>----------</td>
</tr>
<tr>
<td>Face-to-face Course</td>
<td>Spring 2014</td>
<td>Online Course</td>
<td>Undergraduate</td>
<td>84880</td>
<td>76068</td>
<td>89.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Graduate</td>
<td>6423</td>
<td>5583</td>
<td>86.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fall 2013</td>
<td>Online Course</td>
<td>Undergraduate</td>
<td>9192</td>
<td>8118</td>
<td>88.32</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Graduate</td>
<td>1400</td>
<td>1307</td>
<td>93.36</td>
<td></td>
</tr>
<tr>
<td>Face-to-face Course</td>
<td>Spring 2014</td>
<td>Online Course</td>
<td>Undergraduate</td>
<td>73786</td>
<td>68108</td>
<td>92.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Graduate</td>
<td>6373</td>
<td>5622</td>
<td>88.22</td>
<td></td>
</tr>
</tbody>
</table>
REFERENCES


http://dx.doi.org/10.1037/1082-989X.12.2.121

http://dx.doi.org/10.1037/a0014699


Cambridge, MA: Cambridge University Press.


http://dx.doi.org/10.3102/0162373711413814
