Title
A Comparison of Online Instructional Strategies Across Key Student Outcomes

Author
Ingersoll, George

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

A Comparison of Online Instructional Strategies Across Key Student Outcomes

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy in Education

by

George Morin Ingersoll

2015

ABSTRACT OF THE DISSERTATION

A Comparison of Online Instructional

Strategies Across Key Student Outcomes

by

George Morin Ingersoll

Doctor of Philosophy in Education

University of California, Los Angeles, 2015

Professor Richard L. Wagoner, Co-chair

Professor Linda J Sax, Co-chair

The purpose of this study is to compare student outcomes associated with alternate strategies of delivering online instruction in higher education. More so than traditional classroom education, different strategies for online education may vary widely in format and approach. To date, most research pertaining to online education has focused on comparisons between online and traditional classroom instruction, with few studies dedicated to comparing different strategies of online education. This makes it difficult for educators to effectively choose between the many diverse options for structuring their online programs and courses. This study begins the process of addressing this issue by investigating whether different strategies of online instruction may have variable effects upon several key student outcomes.

To this end, this study conducted an experiment using a single lesson in which the strategy of instruction varied across treatment groups while other key variables: the instructor, content, materials, and timeframe were all held constant. The experiment was conducted with the help of 425 undergraduate student volunteers who were randomly assigned among five treatment groups. Each group was exposed to a different instructional strategy: (1) traditional classroom instruction for control, (2) hybrid, or blended, instruction (3) synchronous instruction through videoconference, (4) asynchronous instruction with pre-recorded video, and (5) asynchronous instruction with text and slides. The effects of these different instructional strategies were evaluated using measures of the participants' comprehension, engagement, satisfaction, and lesson completion rates.

In terms of comprehension and satisfaction, the classroom and hybrid groups had the highest average scores. Conversely, the two asynchronous online groups scored highest on rates of participation, comment quality scores, and lesson completion. The synchronous online group scored the lowest on the assessment of comprehension and on the average quality of discussion comments, but was significantly above the two asynchronous groups on several measures of satisfaction. These variable results suggest that there is no one "best" instructional strategy for maximizing all student outcomes, but that differences in the instructional strategies' effects upon student outcomes do exist. Hence, the most effective strategy depends heavily on context and upon which student outcomes the educator is seeking to maximize.

The dissertation of George Morin Ingersoll is approved.

Mark Kevin Eagan

Megan Loef Franke

Carla Hayn

Richard L. Wagoner, Committee Co-chair

Linda J Sax, Committee Co-chair

University of California, Los Angeles

2015

Table of Contents

Abstract ii

List of Tables viii

Vita ix

Chapter 1: Introduction 1

The Current State of Online Higher Education 1

Problem Statement 2

Background 3

Growth of Online Education 3

Different Online Instructional Strategies 5

Forces Shaping Online and Higher Education 7

How Forces Influence Adoption of Instructional Strategies 10

Knowledge Gap: Why Information About Instructional Strategies is Needed 11

Purpose of the Study and Research Questions 12

Audience and Intended Outcomes 13

Delimitations 13

Limitations 16

Definitions of Terms 17

Chapter Summary 18

Chapter 2: Literature Review 20

Organization of the Chapter 20

Context: The Study of Distance and Online Education 20

Pedagogical Underpinnings: Theories of Learning and Instruction 22

Application of Learning Theories to Distance Education 24

Challenges and Opportunities in Evaluating Online Instructional Strategies 26

Studies of Online Instruction 28

Comparisons of Face-to-Face versus Online Instruction 29

Evaluation of Online Instruction for Specific Objectives 32

Results of Online Instruction by Subject Area 33

How to Implement Online Instruction 33

Outcomes for Different Student Groups in Online Instruction 35

Comparisons of Online Strategies for Instruction 35

Chapter Summary 37

Chapter 3: Methods 40

Introduction 40

Organization of the Chapter 40

Restatement of the Study's Purpose, Significance, and Research Question 40

Introductory Summary of the Methodology 41

Pilot Study 42

Experimental Research Design 43
Discussion of Experimental Design 43
Application of Experimental Design to the Study 43
Online Instructional Strategies as Treatments 45
Group C: Classroom Control (Face-to-Face) 45
Group S: Synchronous Videoconference 45
Group H: Hybrid 47
Group V: Asynchronous Video and Forums 48
Group T: Asynchronous Text and Images with Forums 49
Cross-Group Similarities and Differences 49
Outcome Measures 51
Assessment of Comprehension 51
Survey of Satisfaction 52
Quality and Frequency of Engagement Measured by Discussion Responses 54
Participation/Attendance 55
Population and Sample 56
Chronology of Data Collection Procedures 58
Data Analysis 59
Independent Variables 59
Dependent Variables 61
Statistical Techniques 62
Threats to Validity 63
Chapter Summary 66
Chapter 4: Results 67
Organization of the Chapter 67
Introduction 67
Execution of the Study 68
Demographic Information 72
Overview of Outcome Measures 75
Comprehension 79
Class Participation 81
Likelihood of Participation 81
Comment Quality 83
Satisfaction 85
Full Survey Average 86
Subject Matter Satisfaction 87
Satisfaction with the Instructor 88
Format of the Lesson 90
Opportunities for Interaction 92
Learning Experience 93
Study Completion/Attrition 95

Regression Results 97
Interaction Effects 101
Chapter Summary 108
Chapter 5: Discussion and Implications 111
Introduction and Study Synopsis 111
Purpose of the Study 111
Prior Studies 112
Investigative Approach and Analysis 114
Summary of Findings 116
Discussion of Findings 117
Overview of Notable Observations 117
Quiz Results 119
Engagement 122
Satisfaction 126
Study Completion 131
Interaction Effects 133
Implications for Practice 134
Limitations 139
Potential for Future Research 143
Conclusion 147
Appendices 149
References 168

List of Tables

Table 3.1. Demographic Independent Variables 60
Table 3.2. Key Dependent Variables 61
Table 4.1. Demographic Information of Study Registrants and Participants 73
Table 4.2. Descriptive Statistics of Student Outcomes by Treatment 77
Table 4.3. Tests of Normality 78
Table 4.4. Quiz Score by Group 79
Table 4.5. ANOVA for Quiz Scores 80
Table 4.6. Participation Rates 82
Table 4.7. Chi-Square Tests for Participation Rates 83
Table 4.8. Comment Quality 84
Table 4.9. ANOVA for Comment Scores 84
Table 4.10. Full Survey Averages 86
Table 4.11. ANOVA for Survey of Satisfaction 86
Table 4.12. Survey Subject Averages 87
Table 4.13. ANOVA for Subject Questions 88
Table 4.14. Survey Instructor Averages 89
Table 4.15. ANOVA for Instructor Questions 89
Table 4.16. Survey Format Averages 91
Table 4.17. ANOVA for Format Questions 91
Table 4.18. Survey Instructor Averages 92
Table 4.19. ANOVA for Interaction Questions 93
Table 4.20. Survey Learning Experience Averages 94
Table 4.21. ANOVA for Learning Experience Questions 94
Table 4.22. Study Completion 95
Table 4.23. Chi-Square Tests for Study Completion Rates 96
Table 4.24. Significant Coefficients from all Regressions at 95% Confidence 99
Table 4.25. Interactions for Quiz Score between Age & Treatment Group 103
Table 4.26. Interactions for Full Survey between Language & Treatment 105
Table 4.27. Interactions for Questions about Subject between GPA & Treatment 107
Table 4.28. All ANOVA and Chi-Squared Results at 95% Confidence 109

VITA

George Morin Ingersoll is the Director of Hybrid Learning Initiatives at the UCLA Anderson School of Management. At Anderson, George works with faculty to expand the use of media and instructional technology across the curriculum. He is also responsible for spearheading the launch of FEMBA Flex, UCLA Anderson's hybrid scheduling option for part-time MBA students. George has worked in online education since 2004.

PRIOR EDUCATION

UCLA ANDERSON SCHOOL OF MANAGEMENT Los Angeles, CA
MBA (full-time) September 2007 - June 2009

- *Dean's List, Faculty Scholar, Honor Society*

UNIVERSITY OF SOUTHERN CALIFORNIA Los Angeles, CA
B.S. Business Administration September 1997 - May 2001

- *Trustee Scholar (full tuition), National Merit Scholar*

RELEVANT WORK EXPERIENCE

UCLA ANDERSON SCHOOL OF MANAGEMENT Los Angeles, CA
Director of Hybrid Learning Initiatives June 2011 - Present

- Designed and launched UCLA Anderson's first blended-learning MBA curriculum in September 2012.
- Implemented new online learning strategies, including classroom and studio recording, videoconferencing, discussion forums, dropboxes, and online assessment tools across all MBA programs.
- Collaborated on marketing/enrollment that increased the size of the part-time MBA class by 11%
- Developed online course preparation and support workshops for all incoming MBA students.
- Teaching Fellow for Statistics, Accounting, Negotiations, and Entrepreneurship

ICDC COLLEGE Culver City, CA
Online Campus Director/ Director of Education April 2009 - May 2011

- Launched ICDC College's first Online Campus with three Associate Degree programs in August 2009. Grew the campus to over 1500 active students in less than two years.
- Supervised all departments serving the Online Campus. From two employees at launch, hired and trained a staff of over 150 full-time employees and adjunct faculty.
- Managed all operational functions of the Online Campus, including the launch of new programs, IT implementations, development of employee reporting and evaluation systems, and other projects.
- Oversaw a successful first accreditation visit from ACCSC.

LAUREATE ONLINE EDUCATION/ WALDEN UNIVERSITY Baltimore, MD
Service and Operations Manager August 2006 - May 2007
Senior Course Manager September 2005-August 2006
Course Manager September 2004-September 2005

- Supervised 21 managers and specialists of the Baltimore Course Management & Financial Services Teams.
- Managed the design, testing, and training of the company's new CRM System for all front-line support.
- Consolidated customer service within one department and expanded coverage from 5 to 7 days per week.
- Directed customer support for Walden's successful transition to a new learning management system.
- Analyzed Walden's classroom set-up process, and used the resulting information to reengineer the set-up procedures for courses serving ~40,000 students, reducing classroom set-up times by more than half.

OTHER WORK EXPERIENCE

Primary School Librarian, Teach for America Baltimore, MD
LSAT, GRE, MCAT, & SAT Instructor, Kaplan Test Prep Baltimore, MD
SAT Instructor, Sylvan Learning Systems Los Angeles, CA

Chapter 1: Introduction

The Current State of Online Higher Education

Enrollment in online classes has grown by over 350% in the U.S. since 2003 and, with over one third of all U.S. college students taking at least one online course in 2013, the evidence suggests that online education has become, and will continue to be, a major component of U.S. higher education (Allen & Seaman, 2014). While once primarily the province of for-profit and vocational schools, online education has rapidly gained acceptance among traditional universities (Christensen & Eyring, 2011). In a 2013 poll, over 90% of college presidents described online education as a critical component of their schools' long-term strategies, indicating that, in spite of the major proliferation of online classes in recent years, many schools are still developing strategies for online education (Allen & Seaman, 2014).

This rapid growth in online education means that faculty and administrators face a variety of up-front decisions about how to implement and structure their online programs. These are important and challenging decisions because, unlike traditional classroom teaching, strategies for online education may vary widely from one program to the next (Kearsley, 1997). In online classes, choices about what software, media, and collaboration tools to implement combine with operational decisions such as whether to offer online instruction live (synchronously) or using prepared content (asynchronously)—and whether to include any face-to-face components.

Unfortunately, college administrators and faculty have little empirical data to use when making critical decisions about which of these online instructional strategies to adopt, and any understanding of these different strategies' effects upon key student outcomes is limited at

best (Maguire, 2005; Parthasarathy & Smith, 2009). Because strategies for online education vary widely in approach, it is problematic to compare online instruction across schools, programs, or even instructors. One online class might be driven primarily by asynchronous forums (online discussion threads that evolve over several days and do not require simultaneous participation) whereas another could be based on synchronous video interactions (live video chat or webinars). Still other online classes may be built around adaptive assessments (interactive tests that adjust later questions and instructional content based on each student's prior answers) or even simple text-based lessons (similar to correspondence courses). Because research on online education is still in its infancy, there is little evidence indicating which of these strategies may be more or less effective in improving key student outcomes such as comprehension, satisfaction, and engagement.

Problem Statement

This lack of research comparing the effectiveness of different strategies for online education delivery makes it difficult for administrators and instructors to effectively choose between the many options for structuring their online education programs and courses. In the absence of this knowledge, decision-makers are essentially taking a "best guess" approach as to how best to achieve their desired learning outcomes. Also, because the implementation of online education programs usually requires a significant fixed investment in terms of both funding and effort, it is often costly and difficult to change the mode of online education delivery once it has been implemented (Ng, 2000; Boettcher, 2004). The importance of making sound decisions up-front is magnified by the fact that much of the technology used in the delivery of online education requires schools to sign fixed-term contracts for the use of

software and/or services in addition to possibly having to hire staff to support the new mode of instruction (Rumble, 2001). Thus, once online educators are "locked-in" to one particular strategy, they may have significantly less leeway to adjust their mode of delivery as compared to instructors teaching in a traditional classroom setting. This makes it all the more important that decision-makers have sound, empirical results on the effectiveness of different forms of online education so that they can make informed decisions when developing courses.

Background

Growth of Online Education

Globally, the demand for higher education is growing rapidly despite a critical lack of physical infrastructure to support so many new students (Howell, Williams, & Lindsay, 2003). In the U.S., the Department of Education's National Center for Education Statistics (NCES) (2012) predicts further growth in higher education enrollment of between thirteen and fourteen percent by 2021. Meanwhile, approximately one billion new users worldwide have gained internet access between 2008 and 2012 (Meeker & Wu, 2013). Given this significant growth, it is understandable that so many universities have turned to online education as a means for addressing this anticipated shortfall in available classroom seats. The response has been strong: enrollments in online classes at the post-secondary level have risen at a compound annual growth rate of 16.2% per year between 2002 and 2012 in the U.S., as compared to 2.5% for higher education as a whole (Allen & Seaman, 2014). The global appetite for U.S. college courses delivered in an online format has also been impressive: Coursera, one of the leading platforms for Massive Open Online Courses (MOOCs), announced in 2012 that it had recorded over one million course enrollments within just four months of its launch (Coursera Blog, 2012).

Reflecting the growing popularity of online courses, in its most recent report on Higher Education, the New Media Consortium (Johnson, S., Estrada, & Freeman, 2014) lists the integration of online learning as one of two most important factors driving change in higher education over the next two years.

Of course, the concepts of providing education at a distance and/or through the use of instructional technology are by no means new. Distance education, which dates back as far as the 1840s in England (Matthews, 1999), can be viewed as the evolutionary ancestor of modern-day online instruction (Bernard, et al., 2004; Larreamendy-Joerns & Leinhardt, 2006). Simonson (2003) defines distance education as "…institution-based, formal education where the learning group is separated, and where interactive telecommunications systems are used to connect learners, resources, and instructors." With the inclusion of the internet, this definition could as easily be applied to online education as well. To understand the development of distance education, Taylor (2001) outlined five "generations" of distance learning beginning with print correspondence, then the inclusion of other forms of instructional media, followed by synchronous communication through audio and videoconferencing, then the use of interactive multimedia and computer-mediated communication, and finally incorporating all institutional resources and processes through a comprehensive online portal. More recently, Keengwe and Kidd (2010) have offered a similar framework for viewing the development of online education. Their "eras" of online instruction include Computer-Assisted Learning from 1975-85, Computer-Based Training Using Multimedia from 1983-1990, Web-Based Education and Training from 1990-1995, eLearning from 1995-2005, and Mobile Learning and Social Networking from 2005 and on.

It is important to note that in both the case of Taylor's "generations" as well as Keengwe and Kidd's "eras," all earlier forms of distance and online instruction are still in use today. Hence, while some schools are right now implementing the latest forms of adaptive assessments using mobile devices, others continue to operate correspondence courses by mail (Parsad & Lewis, 2008). As internet-based technologies continue to evolve at a fast pace, and with the rapid adoption of MOOCs to serve as an example, there is no reason to believe that these trends towards the development of more numerous and diverse strategic options for online and distance education are likely to diminish in the near term (Billington & Fronmueller, 2013; Hardesty, 2013). Even making sense of the myriad forms that online instruction can take now is a formidable task in itself.

Different Online Instructional Strategies

At the broadest level, Allen and Seaman (2014) define "Online" courses as delivering 80% or more of their content online, "Hybrid or Blended" courses as being made up of between 30% and 79% online content, and any other course with some smaller amount of online content as "Web-Facilitated." These are useful definitions, but they hardly do justice to the myriad strategies that may exist within any of these broad categories. As Bowen et al (2012, p. 7) point out in their report on online learning:

> …online education is hardly one thing. It comes in a dizzying variety of flavors, ranging from simply videotaping lectures and posting them for any-time access, to uploading materials such as syllabi, homework assignments, and tests to the Internet, all the way to highly sophisticated interactive learning systems that use cognitive tutors and take advantage of multiple feedback loops.

To extend this point, because of the many ways that different tools, techniques, software, and formats can be combined into a single strategy for online instruction, it would not be a far stretch to categorize the online learning strategies of most institutions as entirely unique.

Nevertheless, it is possible to identify some critical differences between strategies which will provide more insight into overall instructional approaches than the categories provided by Allen and Seaman. These differences include the critical distinction as to whether instruction is delivered synchronously or asynchronously. Synchronous online instruction involves groups of students and the instructor all being online at the same time and communicating in real time. Asynchronous instruction involves a delay in communication and may feature students interacting with instructional content which was prepared by the instructor in advance in order to be used by students on their own schedules. Among the two formats, asynchronous is considerably more prevalent: The National Center for Education Statistics (2008) reports that among all post-secondary institutions which offer some distance education courses, 92% use asynchronous instruction to a moderate or large extent, as compared to 31% for synchronous. In addition to timing, one more important distinction among online education strategies is the type of media which is employed in instruction. Among the possibilities are video, audio, text, slides, adaptive assessments, simulations, games, and more. Many online courses utilize several different types of media for instruction, but video and text are the most popular (Parsad & Lewis, 2008).

Another differentiating factor among courses which utilize online instruction to a large extent is whether the course includes any time spent in the classroom. As noted above, this type of strategy is defined by Allen and Seaman as hybrid or blended instruction. In hybrid

courses, the online content may be synchronous or asynchronous, but is more often the latter (Helms, 2012; Parsad & Lewis, 2008). Specifically within this category of hybrid courses which utilize asynchronous online content, there is an instructional technique known as "flipping the classroom" that has recently gained in popularity (Berrett, 2012). Bishop and Verleger (2013) define the flipped classroom as "an educational technique that consists of two parts: interactive group learning activities inside the classroom, and direct computer-based individual instruction outside the classroom." Flipping the classroom, while not alone among popular new online instructional strategies, stands out for having received high-profile and favorable press from leading publications such as the New York Times (Fitzpatrick, 2012) and Science (Mazur, 2009). This recent surge in popularity for flipping the classroom and other cutting-edge online strategies raises the question of what factors in higher education are driving the growth of online education and influencing the adoption of specific strategies.

Forces Shaping Online and Higher Education

In their report on distance education at the post-secondary level, The National Center for Education Statistics (Parsad & Lewis, 2008) lists the results of a survey of higher education institutions regarding the factors that influenced their decisions to utilize online instruction. The factors which were most frequently cited as affecting to a major extent the schools' decisions to launch online courses were as follows (p. 16):

- Meeting student demand for flexible schedules (68%)
- Providing access to students who otherwise would not have access (67%)
- Making more courses available (46%)
- Seeking to increase student enrollment (45%)

This data seems to indicate that the most important factors for many schools adopting online education strategies have to do with attracting and retaining more students—particularly those students who might not otherwise have applied to a traditional program. This position is supported by data which indicates that a majority of post-secondary students who are studying online would not have enrolled in a traditional program if an online option had not been available (Aslanian & Clinefelter, 2013; Hrastinski, 2008b) The reason for this appears to be the high value that online learners place on the scheduling flexibility afforded by many online and hybrid programs of study (Ausburn, 2004). As Aslanian and Clinefelter (2013, p. 5) note, "The greatest advantage of online study continues to be scheduling flexibility and the freedom to manage other responsibilities."

Another force which represents an important consideration for most higher education institutions is cost (Ehrenberg, 2012), and there are a fair number of studies that have analyzed the expenditures of monetary and time associated with online instruction. One on side of the issue of costs are studies such as Boettcher (2004), Ng (2000), and Rumble (2001) which make the case that online courses cost more in terms of money and development time than traditional instruction. These studies typically cite higher up-front costs to build and launch online courses. As a counterpoint, other researchers such as Benoit, Benoit, Milyo, and Hansen (2006) plus Vilaseca and Castillo (2008) argue that the long-term savings afforded by reusable instructional content makes online education a more cost-effective option over time when compared with classroom instruction. Still others point out that the costs of online instruction can be more than offset by increased tuition revenues from additional enrollments, and suggest that online courses may be used to offset the decline in state support at many public

universities (Byrd & Mixon, 2012). What is highly apparent in all of these analyses of costs, however, is that the strategy of online instruction which is employed plays a critical factor. According to the calculations of one study by Arizona Learning Systems (1998), the costs associated with the development of one online course may differ by a factor of ten or more according to the type of online instructional strategy being used.

Finally, some academics such as Christensen and Eyring (2011) have granted an air of inevitability to the continuing growth of online education and its potential to supplant large amounts of classroom instruction. Noting the low variable costs of asynchronous online instruction as well as its ability to scale up quickly to accommodate large numbers of students, they argue that online education fits all the standard criteria for disruptive innovation as identified by Christensen in the business world (Christensen C. , 1997). Pointing to rising tuitions at traditional universities while MOOCs may offer the same quality of instruction for free, Christensen states, "Higher education is just on the edge of the crevasse. Generally, universities are doing very well financially, so they do not feel from the data that their world is going to collapse. But I think even five years from now these enterprises are going to be in real trouble." (Howe, 2013) Other writers have hurried to compare the state of traditional higher education to the music, newspaper, and postal service industries at the turn of the millennium—each of which was significantly disrupted by the internet (Keller, 2011; Carey, 2009). Regardless of whether or to what extent these predictions are accurate, or to what extent one agrees with them, many of these viewpoints are shaping not only the overall growth of online higher education, but also the types of online instructional strategies that are being adopted.

How Forces Influence Adoption of Instructional Strategies

As previously noted, evidence from the National Center for Education Statistics suggests that a primary motivation for many schools to build online courses is to attract new students by offering, among other things, flexible schedules of instruction. Sadly, there is a considerable amount of evidence that persistence rates in online courses are lower (Brady, 2001; Carnoy, Rabling, Castano-Munoz, Montoliu, & Sancho-Vinuesa, 2012; Diaz, 2002; Gleason, 2004). Since increasing student enrollment appears to be the goal of many schools with online learning programs and higher drop rates are detrimental to that outcome, it makes sense that many of these schools would seek to adopt strategies for online instruction that lead to the highest rates of persistence. Unfortunately, the research does not currently identify which, if any, online instructional strategies outperform the others in terms of retention rates. If such a strategy could be identified, then it presumably might be an attractive option for schools seeking to grow enrollment.

Another important factor which may influence the type of strategy that a school adopts is the extent to which the strategy adheres to traditional forms of successful classroom teaching. Perhaps partly because online education is relatively new, studies have shown that online education still has something of a perception problem with the U.S. public (Saad, Busteed, & Ogisi, 2013). At the very least, online is not regarded by everyone as having equivalent outcomes to traditional education (Allen & Seaman, 2014) and concerns about academic dishonesty in online courses persist (Khare & Lam, 2008; Cole, Shelley, & Swartz, 2012). Because of these negative perceptions, it may be guessed that schools will be more likely to gravitate towards online instructional methods which are more familiar and represent

the smallest departure from traditional classroom teaching. These relatively "safer" strategies include hybrid instruction and some synchronous online strategies, since synchronous strategies such as videoconferencing are often perceived as most similar to face-to-face teaching (Lieblein, 2000). On the other hand, there are some enthusiastic educators who strongly believe that outcomes for online courses can exceed classroom instruction – especially if carried out differently. Educators who fall into this category are most likely to embrace asynchronous online strategies (McDonald, 2002). MOOCs also, with their departure from traditional instruction and their need to scale enrollments at levels far beyond any student-to-instructor ratios that would be possible in a classroom setting, are also more likely to make use of asynchronous instruction (Skiba, 2012).

Knowledge Gap: Why Information About Instructional Strategies is Needed

As has been cataloged above, online education is hardly a single strategy for the delivery of instruction. Rather, it consists of myriad approaches, each of which may have distinct advantages and drawbacks. And, while there are many pressures within higher education that are pushing the growth of online education to outpace enrollments in traditional classes (Allen & Seaman, 2014), it is still unclear which online education strategies will work best to address those pressures. Some issues, such as cost (Rumble, 2001) and scalability (Skiba, 2012), can be more easily measured across strategies, but, in terms of other important factors, particularly student learning outcomes, engagement, and satisfaction, it is still unclear which strategies are most effective. As more schools and departments look to build online classes, educators and administrators need as much information as possible in order to effectively design and implement online learning programs that will meet their students' needs. Lacking knowledge

about how different strategies may affect critical student outcomes, educators may gravitate towards strategies that use familiar forms of instruction (Lieblein, 2000) or have other easily-measured benefits such as cost-savings, when these strategies may not necessarily be the best for student learning or satisfaction (McDonald, 2002). Empirical evidence which directly compares online instructional strategies across student outcomes is needed for these educators to structure their online courses most effectively.

Purpose of the Study and Research Questions

The purpose of this study is to look for evidence of whether there may be differences between several major strategies of online instruction in terms of affecting the key student outcomes of comprehension, satisfaction, participation, and engagement. To date, most of the studies addressing online education at the post-secondary level have assumed a somewhat monolithic view of online instruction when, in reality, there are a wide range of very different strategies that can be employed. Through the use of an experiment in which the strategy of online instruction varies but the content of the lesson (aside from aspects of its delivery) does not, this study seeks to begin the process of addressing the following broad research question: "Does the type of online instructional strategy employed have an effect upon key student outcomes?" More specifically, "Are there differences across students' comprehension, satisfaction, participation, and engagement according to whether a lesson is taught in class, in a hybrid format, online using videoconferencing, online using pre-recorded video, or online using text and images?" By providing initial evidence which may link the delivery strategies of online instruction to important outcome measures, the goal of this study is to provide results which may help practitioners to make better informed choices as they develop their online lessons.

Audience and Intended Outcomes

The intended audience for this study is higher education instructors and administrators who are involved in the creation of online instructional content. This may include everyone from experts in online instructional design to novice faculty who are developing their first courses using online content. The hope is that, by reading this study, these educators may gain insight into how different online instructional strategies may affect important student outcomes. Combining this information with other key considerations such as cost, ease of use, and limitations on resources, ideally these educators will be able to make better-informed decisions about which online instructional strategies will best fit the needs of their students and institutions.

Delimitations

This study compares student outcomes across online instructional strategies for only a single lesson—as opposed to an entire course or program of study. A study with a longer timeframe would provide additional useful information, particularly about persistence rates, but would lose some of the focus on specific learning outcomes. Because the objective of this study is to examine how students engage with and perceive a lesson according to its method of delivery, evaluating only a single lesson allows for greater insight into more narrowly-focused learning outcomes and avoids interfering factors that might come up over an entire course or program. It would also be very difficult to conduct a true experiment across an entire multi-unit course being offered in several different formats because participants would most likely want to have some say in the type of instruction they would receive.

In addition, although the topic of how students with various learning styles may engage differently with online instructional content is interesting and certainly worth exploring, this study does not introduce learning styles into its analysis. Although demographic information about participants was captured in a pre-study questionnaire (see Appendix A), it would have been unreasonable to ask student participants to identify their own learning styles and expect anything approaching total accuracy (Dunn, 1983). Also, categorical data about the participants was not used to assign the participants to treatment groups because to do justice to those comparisons would have required a different experimental design that would not have allowed for the types of broad comparisons between instructional strategies that this study has pursued.

In a similar vein, this study drew its entire sample group from the population of undergraduate students at a single, selective public research university. While drawing participants from a broader student spectrum might have provided a more representative sample of the entire higher education population, to do so would have diminished the homogeneity of the treatment groups and might have resulted in less-reliable comparisons between instructional strategies.

For the same reason, this study did not examine the teaching of multiple subjects using online education. While it makes sense that some subjects might be more or less effectively taught using certain instructional strategies, it would have been more difficult to compare the effectiveness of the instructional strategies themselves if more subjects were included. Instead, the lesson utilized by this study was a subject of business law that requires both comprehension and application. Because of this choice of lesson, the student outcomes of

comprehension and engagement were able to be evaluated without additional confounding factors.

Also, as previously noted, strategies of online instruction can take on a large variety of forms, not all of which could have been realistically evaluated in a single study. While some newer strategies such as adaptive assessments, virtual reality, and blends of synchronous and asynchronous online instruction (Power, 2008) are all potentially interesting strategies to explore, this study focused instead upon some of the most common and clearly-defined methods of online instruction. In spite of this focus, however, there exist major differences in approach even within more narrowly-defined strategies. For example, time and production costs spent in creating online content and platforms may have the potential to influence both the real and perceived effectiveness of online instruction. In this study, all instructional techniques utilized fairly basic and inexpensive equipment and software—as may be found in most typical U.S. university campuses. In short, this study made no distinctions drawn between the effectiveness of different levels of production quality within strategies of online instruction, though that may certainly be a potentially interesting avenue of future research.

Finally, the review of the literature for this study does not delve deeply into studies having to do with offline distance education. While there is ample research dealing with instructional strategies in distance education, these strategies are quite different from modern online instructional strategies. Other than the most general comparisons between distance education and classroom instruction, the research on distance education does not provide a great deal of insight into the effectiveness of strategies for online education. Therefore, this study addresses only the prominent general theories having to do with distance education and

does not discuss a majority of the research which deals with specific offline distance education strategies.

Limitations

Many of the most notable limitations of this study have to do with the above delimitations. For one thing, as this study has focused on a single lesson, it may not be practical to extend the conclusions drawn from this single lesson to an entire course. It could be that, over several lessons, differences in outcomes between instructional strategies could become more or less pronounced. Therefore, decision makers may wish to consider additional evidence when deciding how to implement a full online course or an entire program of study. Nor will these same results necessarily be applicable across all subject areas. Depending on the type of information being conveyed or the level of interaction and higher-order thinking required in a lesson, the most effective form of online instructional strategy may change.

Additionally, the results from the students who participated in this experiment may not necessarily be representative of all undergraduate students. This study strove to maximize homogeneity among participants to help control for differences between the treatment groups, but it is certainly possible that some groups of students may not have the same reaction to different instructional strategies than this group of students from a public research university. Students who are significantly less tech-savvy, for instance, or who have greater constraints upon their time, might have very different experiences with classes taught in various instructional formats.

Definitions of Terms

- "Distance Education"/ "Distance Learning" – any form of instruction in which students do not meet in person. Includes online education.
- "Online Education"/ Online Learning" – any form of distance education in which instruction is delivered using the internet.
- "Traditional"/ "Classroom"/ "Face-to-Face Education" – Instruction delivered by the instructor in person to a physically-present group of students, usually in a classroom setting.
- "Hybrid Learning"/ "Blended Learning" – Courses taught using a combination of online and classroom instruction.
- "Synchronous Online Instruction" – Online instruction that occurs at a scheduled time of day with the instructor and students connected live via the internet and communicating in real-time.
- "Asynchronous Online Instruction" – Online instruction in which communication between course participants is intermittent and unscheduled. Asynchronous courses are not necessarily self-paced – they may have deadlines, but do not involve real-time, live interaction.
- Massive Open Online Course (MOOC) – a free online course that is open to all participants. Typically boasts enrollment numbers that are much larger than could be contained by almost any classroom—in some cases many thousands.

Chapter Summary

Online education is still in the early phase of its development. As Phillip Schmidt of Peer 2 Peer University recently stated, "The field is now at the stage where film was when the first movie cameras became available and people immediately mounted them at the backs of theaters to record stage plays." (Hardesty, 2013) While there are several powerful forces (demand, efficiency, flexibility, etc.) that are driving the growth of online learning, these forces are not necessarily shaping the development of the field in similar ways. As a result, instead of taking on a single, easily-identifiable form, online education strategies have diversified. Some of the most important distinctions between online strategies include whether the instructional content is delivered synchronously or asynchronously, the type of media employed in instruction, and whether there is a classroom component to the instruction. What is poorly understood about these different strategies is whether there are any differences between them in terms of how they affect key student outcomes such as learning, satisfaction, and engagement. Without this critical information, educators will have difficulty in making optimal decisions about which online instructional strategies to adopt so as to most-effectively serve their students and institutions. This study seeks to shed additional light on this decision-making process by presenting the results of an experiment in which a single lesson was delivered to several groups of students—each using a different online instructional strategy. Each treatment group was then evaluated for comprehension, satisfaction, engagement, and participation, with the results compared across the type of instructional strategy employed. The hope is that the results of this experiment may be used by educators who are building

online content to make better-informed choices about which online instructional strategies to employ in order to maximize the student outcomes that they most value.

Chapter 2: Literature Review

Organization of the Chapter

This chapter provides an overview of the literature which is most applicable to the study of online education. The first section sets the stage by introducing a prominent theory of distance education which refutes the idea that differences exist across different instructional delivery mechanisms and media. The next section provides a broad analysis of traditional theories of learning and instruction, followed by a discussion of the ways in which these traditional theories can be applied to modern online education. The subsequent section details some of the challenges and opportunities inherent in the study of online education, which then leads to an overview of the existing studies dealing with online education. This review of the research having to do with online education is broken down into several categories, culminating with the research that is most similar and directly relevant to this study.

Context: The Study of Distance and Online Education

Thirty years ago, Richard Clark (1983) argued that the type of media employed in instruction has no effect whatsoever upon student learning outcomes. Rather, he stated, it is only the quality of the instruction itself that matters. Clark employed the clever analogy of a truck used to deliver groceries to equate media to a delivery vehicle; making the point that no matter what type of vehicle is used, it will not alter the nutritional content of the food delivered. Likewise, he argued, whether instruction is delivered in person, by correspondence, or via any other media or device, the students' learning outcomes will remain the same. Naturally, this viewpoint has hardly been without its critics, and later writers responded to Clark's claim by pointing out that some approaches to instruction are only made possible by the

use of certain types of instructional media (Reiser, 1994; Kozma, 1994). Of course, these writers were debating the impact of instructional media well before the widespread adoption of the internet and major proliferations of revolutionary instructional technology such as learning management systems (Sheridan, White, & Gardner, 2002), so perhaps a more up-to-date reexamination is in order.

Modern online instruction can be characterized by a number of important attributes. With regards to communication in the course, it can be asynchronous or synchronous, one or two-way, and may be between the student and the instructor, small groups, the full class, or, in the case of Massive Open Online Courses (MOOCs), between thousands of learners. Instruction, assessment, and feedback in an online course may be given in real time or delayed, and may be auto-programmed and generic or personalized to each student. The choice of media for instruction is equally diverse, with video, audio, text, adaptive assessments, games, simulations, and other interactive activities all employed (Ely, 2003). At the more granular level, the many options in terms of software, learning management systems, and content delivery mechanisms may all potentially complicate decisions faced by faculty and administrators who are responsible for building online courses. Finally, practitioners must now make decisions about which devices (laptops, tablets, mobile phones) to target with their educational content and potentially even which operating systems to develop for. This study takes the view that, in light of the many new developments in internet-enabled instructional technology, Clark's argument deserves at least another look.

Unfortunately, research into the field of online education is naturally in its early stages since public use of the internet on anything approaching a large scale is still less than thirty

years old (Leiner, et al., 2009). The first Learning Management System was launched less than twenty years ago by the University of Auckland School of Management in 1995 (Sheridan, White, & Gardner, 2002) and by 2002 still less than 10% of higher education students were taking classes online (Allen & Seaman, 2014). Given this relatively short timeframe, it is understandable that researchers are still in the process of developing a fundamental understanding of what works and what does not in online education. Nevertheless, online education has its roots in distance education (Bernard, et al., 2004), to which some considerable attention has been devoted, and, even more fundamentally, in the general theories and research having to do with traditional forms of learning and instruction (Anderson, 2008; Conrad & Donaldson, 2011). Because this study seeks to understand how the use of different strategies for online learning may affect certain critical student outcomes, and this issue is, at its core, a question of curriculum and instruction, before Clark's theory can be conclusively addressed, it is necessary to provide some broader context regarding educational theory by delving into the influential research and writing on traditional learning and instruction.

Pedagogical Underpinnings: Theories of Learning and Instruction

Among the most important modern theories of learning is the theory of constructivism, the principals of which were originally championed in education by Piaget and Dewey in the first half of the 20th Century (Piaget, 1929; Dewey, 1938). Dewey emphasized action in the learning process and believed that learning needs to be experienced in a social setting where ideas and knowledge can be created as part of a community (Dewey, 1938). Piaget also believed in the importance of experiential learning and regarded the ideal role for a teacher as

a facilitator for peer-to-peer interaction and knowledge development (Piaget, 1970). Both of these men, in addition to Bruner (1966), strongly felt that students need new knowledge to be related to their experiences with, and understandings of, more familiar concepts.

Vygotsky (1980), like Piaget, Dewey, and Bruner also espoused the importance of interaction in learning. In contrast to Piaget, however, Vygotsky believed that instruction from a knowledgeable teacher is key to advancing students' understanding beyond what they would be able to learn on their own. Vygotsky defines this gap between what can be learned independently versus with the help of an instructor as the Zone of Proximal Development.

In even more of a departure from the constructivist principles outlined above, Tyler's *Basic Principles of Curriculum and Instruction* (1949) outlines an objectivist approach to education that emphasizes linear curriculum development and an input-process-output model of instruction that stresses the teaching of reality-based facts that are independent of individual interpretation. Tyler also does not shy away from the use of assessment and, in fact, considers it a critical part of the curriculum design process.

More specifically related to the topic of this study, Knowles, in his book, *The Adult Learner: A Neglected Species* (1973), updated the concepts of some of the constructivists, such as Dewey, and interpreted them for adult learners through the concept of andragogy, which deals with instructional strategies geared towards adults. Like many of the other prominent educational psychologists mentioned here, Knowles argues that learners (particularly adult learners) need instructional concepts to be relatable to their lived experiences and benefit particularly from experiential instruction. Knowles also places strong emphasis on the importance of self-direction and intrinsic motivation for the success of adult learners. In a

more-recent update to Knowles' work, Hase and Kenyon (2000) introduced the concept of heutagogy, which expands upon andragogy and emphasizes that adult learners must develop the necessary skills to be able to advance their own educations autonomously.

Application of Learning Theories to Distance Education

All of this, perhaps, raises the question of how to interpret these different viewpoints concerning traditional education in the context of online instruction. Examined from one perspective, the importance that many of these prominent educators and psychologists ascribe to experiential and social learning appears to stand in opposition to the early models of distance education, which typically placed heavy emphasis on the transmission of information from teacher to student (Taylor, 2001). Vrasidas (2000) notes that it was the objectivist ideas of education, as typified by Tyler, which often formed the foundation of most early forms of distance education. Vrasidas also points out that more recent advancements in online instructional technologies, which provide greater opportunities for group communication and collaboration, have made interactive, constructivist approaches to distance education much more viable. As Gunawardena and McIsaac (2003, p. 363) explain, "Technological advances have already begun to blur the distinction between traditional and distance education settings."

Some modern theorists have further argued that the principle of independent learning behind andragogy and heutagogy can be applied with success to online education, particularly when teachers and curriculum designers deviate from the foundational principles of traditional face-to-face instruction and start to utilize strategies for learner empowerment and self-direction that can more easily be applied in an online environment (McDonald, 2002; Eberle &

Childress, 2005). The elements of self-direction and intrinsic motivation, which Knowles emphasizes heavily in andragogy, have also been shown to be of great importance to online learners (Ausburn, 2004). In addition, the constructivist idea that learners gain knowledge through personal interactions with and interpretations of their environments (Smith & Ragan, 1999) makes a certain kind of sense in the context of online education when the environment that the learners are engaged with is the internet and the knowledge it contains (Conrad & Donaldson, 2011). Ultimately, educators who are designing online courses have a variety of classic pedagogical approaches to choose from according to which online instructional strategies they decide to employ. Nevertheless, though technologies now allow educators to create online experiences that are similar to traditional forms of educational delivery (Gunawardena & McIsaac, 2003), this does not mean that all strategies of online education strive for this goal.

In the 1970s, Moore (1993) introduced a pedagogical concept known as Transactional Distance that provides a framework for classifying different strategies of distance education according to levels of student autonomy and the opportunities for direct communication between the students and the instructor that are present. Since its initial introduction, the Theory of Transactional Distance has been updated to keep pace with technological innovations in online instructional delivery (Jung, 2001). Moore was chiefly concerned with understanding how the communication of ideas can break down in an instructional environment when teachers and learners face limitations in their abilities to communicate and collaborate with one another; i.e. at a distance. He therefore defined Transactional Distance as a continuous variable that represents “a psychological and communications space to be crossed, a space of

potential misunderstanding between the inputs of the instructor and those of the learner" (Moore, 1993, p. 22). In Moore's framework, the factors which contribute to Transactional Distance include

- the amount of dialogue between instructors and students,
- the amount of interaction that the students have with each other,
- the structure or rigidity of the course presentation, and
- the level of autonomy that is required of the students in completing course tasks.

All of these factors, taken together, contribute to the amount of Transactional Distance in a course, defined as a continuous variable. Unfortunately, the results of studies that seek to link Transactional Distance to student outcomes have been inconclusive (Gorsky & Caspi, 2005). In fact, evaluating the effectiveness of different forms of online instructional strategies presents considerable challenges across the board.

Challenges and Opportunities in Evaluating Online Instructional Strategies

One major difficulty in making comparisons between online instructional strategies is the relatively high cost of implementation for online courses. Researchers and practitioners have widely documented that the development of an online course typically requires significantly more investment in time and money up front than a traditional classroom course (Rumble, 2001; Boettcher, 2004; Ng, 2000). One early study estimated the cost to develop one 3-unit online course using audio and video as the primary means of instruction at $120,000 (Arizona Learning Systems, 1998), while a more recent study estimated the costs to develop online courses as $25,000 per unit (Boettcher, 2004). Thus, it is exceedingly rare to find a school running two or more sections of a course which employs multiple online instructional

strategies. It is far more common to find schools running traditional face-to-face classes alongside a single type of hybrid or online course. Thus, while it is possible to find many opportunities to compare face-to-face with a single form of online instruction, the number of promising quasi-experiments which could be used to compare different online instructional strategies is extremely limited, at best (Johnson G. M., 2006). Therefore, the fact that the preponderance of research concerning online education has focused on comparing (monolithically-defined) online classes with face-to-face instruction should come as no surprise.

On the other hand, many forms of online education—by their own digital natures—offer ample opportunities for data collection regarding student activities and outcomes. As the latest *New Horizon Report* from the New Media Consortium states, "As learners participate in online activities, they leave an increasingly clear trail of analytics data that can be mined for insights," (Johnson, S., Estrada, & Freeman, 2014, p. 12). In other words, the academic footprints that online students leave are often more accessible and easier to analyze for educators and administrators than those in traditional classroom environments (Siemens, 2011; Campbell, DeBlois, & Oblinger, 2007). A technique known as "educational data mining," which is greatly facilitated by the use of online learning management systems, allows educators to glean insight into broad trends pertaining to instruction and student activity and to apply those insights to specific cases (Romero, Ventura, & Garcia, 2008). For instance, some researchers and administrators have already begun using data mined from their colleges' learning management systems concerning activities that correlate with student success or failure in order to identify and flag students who may be at risk of underperforming (Macfadyen & Dawson, 2010; Fritz, 2011; Morris, Finnegan, & Wu, 2005). In addition, the huge enrollment in MOOCs (Coursera

Blog, 2012) offers rich opportunities for analysis of student activities online in a "big data" environment—particularly regarding those activities which correlate to persistence rates (Ho, et al., 2014; Guo, Kim, & Rubin, 2014). Therefore, while weighing the advantages and disadvantages between different online instructional techniques may require more opportunities for direct comparisons in a quasi-experimental environment at least, the possibilities for insights into the ways that student engage with their learning environments online that may be gleaned from educational data mining and learning analytics are significant and just now starting to be realized (Siemens, 2011; Romero, Ventura, & Garcia, 2008).

Studies of Online Instruction

To date, most of the existing studies dealing with online education can be placed into one of six categories:

1. Studies comparing outcomes for online or hybrid instruction with face-to-face instruction. Examples: (Xu & Jaggars, 2013a; Means, Toyama, Murphy, Bakia, & Jones, 2010; Russell, 1999);
2. Studies which examine the effectiveness of various online instructional strategies in accomplishing specific educational objectives. Examples: (Vonderwell, Liang, & Alderman, 2003; Tu & Corry, 2003; Walsh, et al., 2012);
3. Investigations into whether certain subjects can be taught more or less effectively in an online format. Examples: (Xu & Jaggars, 2013b; Means, Toyama, Murphy, Bakia, & Jones, 2010; Rhoads, Berdan, & Toven-Lindsey, 2013)

4. Research on how to most-effectively implement certain online instructional strategies in order to achieve specific desired outcomes. Examples: (Ngwenya, Annand, & Wang, 2004; Eom, Wen, & Ashill, 2006; Sendag & Odabasi, 2009);
5. Investigations into outcomes for different student groups in online instructional environments. Examples: (Caspi, Chajut, & Saporta, 2008; Drago & Wagner, 2004; Figlio, Rush, & Yin, 2010); and
6. Comparisons of outcomes between different online instructional strategies. Examples: (Choi & Johnson, 2005; Offir, Lev, & Bezalel, 2008; Hrastinski, 2008a).

Comparisons of Face-to-Face versus Online Instruction

By far the most numerous type of research dealing with online education are studies detailing the results of quasi-experiments conducted by instructors who have run side-by-side sections of their classes taught both in-person and online – the latter representing a wide variety of instructional strategies. Representative examples include Benoit et al (2006), Ernst (2008), and Ward (2004). Some educators, such as Figlio, Rush, and Yin (2010), have even conducted true experiments in which the students taking part were randomly assigned to in-person or online classes. Regardless of the methodology used, a great number of these studies have shown no significant differences in terms student learning outcomes between online and face-to-face instruction (Russell, 1999; Means, Toyama, Murphy, Bakia, & Jones, 2010).

Russell's *No Significant Difference Phenomenon* (1999) is an annotated bibliography of many such studies which support Clark's (1983) argument that the media and delivery mechanisms employed in instruction are irrelevant to student performance. In particular, Russell sets out to demonstrate that there are no differences in terms of student success

between classes taught through distance education and those taught using traditional classroom instruction. Russell's work has been influential in convincing institutions to invest in educational technology but has also come under criticism for providing a dearth of information in terms of which distance education strategies may be more or less effective (Lievrouw, 2001). Hence, the *No Significant Difference Phenomenon*, while presenting a compelling case for the adoption of new, cheaper, and more scalable instructional delivery systems, does not provide specific guidance for educators who are selecting from among different instructional strategies.

Similar to Russell, many of the other influential works comparing online and face-to-face instruction are meta-analyses that examine large numbers of related studies. This includes an oft-cited meta-analysis of online and hybrid education published by the U.S. Department of Education (Means, Toyama, Murphy, Bakia, & Jones, 2010) which shows equal or slightly more positive outcomes for online and hybrid education when compared to traditional classroom instruction. Unfortunately, because this meta-analysis includes comparisons using many different strategies for online education all compared to classroom instruction, but no comparisons between online education strategies, there is little indication as to whether there are different outcomes according to the online instructional strategy employed. Thus, while clearly providing support for the decision to create online education programs, Means et al also does not offer much to administrators and faculty who are grappling with the question of which online education strategies to employ.

Other significant meta-analyses in this vein (Sitzmann, Kraiger, Stewart, & Wisher, 2006; Jahng, Krug, & Zhang, 2007; Allen, Bourhis, Burrell, & Mabry, 2002; Byoun & Zoljargal, 2012) offer similar analyses and conclusions, but are similarly limited in their capacity to make

comparisons between different strategies for online instruction. Interestingly, there is evidence that in these large meta-analyses, even when average outcomes are similar between online and face-to-face instruction, there is greater variability in the outcomes for courses that are taught online (Zhao, Lei, Yan, Lai, & Tan, 2005; Bernard, et al., 2004). This seems to suggest that some online instructional strategies may outperform traditional instruction while some may underperform, but does not indicate which of these strategies specifically may be more or less effective.

As a counterpoint to the various meta-analyses showing no significant difference, recent reports by Xu and Jaggars on community college courses in Virginia (2011) and Washington State (2013a) show significantly worse outcomes for persistence and grades among students taking online classes compared to face to face. Xu and Jaggars' studies are not meta-analyses but are very large-scale (>20,000 students) examinations of student performance within a state's community college system. As with the other studies, however, Xu and Jaggars' work does not distinguish between strategies for online instruction within their samples of online classes.

Another potential concerns with many of these comparisons between in-person and online instruction is that, because the majority of studies are quasi-experiments, there is a strong possibility of self-selection bias (Phipps & Merisotis, 1999; Joy & Garcia, 2000). When students are presented with the option to take an online class, it is reasonable to assume that the students who do so either already have some pre-inclination to feel positively about online education or have other reasons that make online instruction an appealing choice for them. Particularly in studies that compare levels of satisfaction, this form of self-selection may

compromise the results of both quasi-experiments and the meta-analyses which use them (Amlie, 2003).

Evaluation of Online Instruction for Specific Objectives

Moving beyond general comparisons of online versus face-to-face instruction with instructional strategies left undefined, a second type of study that deals with online education includes research that examines the effectiveness of specific strategies of online instruction in meeting pre-defined objectives. These studies may be somewhat more useful for faculty and administrators making choices about designing online courses and programs because they isolate one specific online teaching method and evaluate outcomes for that technique, often comparing results with those of traditional classroom instruction. To illustrate with an example, several of these types of studies have addressed the use of asynchronous online discussion forums to interactively engage students with instructional material. Vonderwell, Liang, and Alderman (2003), for instance, show that asynchronous online discussions can be useful for assessing student understanding and are valued by students who use them. Tu and Corry (2003) demonstrate that asynchronous discussions can "promote constructive thinking" while Swan (2002) explains how learning communities can be built with the use of online forums.

These studies certainly provide useful information for practitioners considering specific online tools, but are less useful when there are multiple strategies of online instruction available to accomplish similar goals, as is very often the case. For example, Bullen (1998) and Walsh et al (2012) both offer evidence for the effectiveness of synchronous web conferencing in achieving similar outcomes to those that Tu & Corry achieved using asynchronous

communication tools. Therefore, practitioners have evidence that both techniques (asynchronous discussions and synchronous web conferencing) can effectively promote critical thinking, but have no way to accurately compare the effectiveness of these two techniques with each other.

Results of Online Instruction by Subject Area

Related to the above research are studies which examine the effectiveness of online instruction for teaching specific subjects. While Means et al (2010) indicates no differences in comparisons of online versus face-to-face instruction by the subject matter being taught, other studies dispute this result. Xu and Jaggars (2013b), for instance, provide evidence that students in online classes perform comparatively worse in English and Social Sciences. Other studies cite the students themselves as expressing the feeling that some subjects do not offer equal opportunities for interaction when taught online (Al-Shalchi, 2009). In addition, other researchers, such as Rhoads, Berdan, and Toven-Lindsey (2013) note the preference of many schools to convert classes in certain subjects ("natural, hard, or applied sciences" (p. 91) to an online format. This may imply that faculty or administrators feel that subjects which focus more heavily on fact-based instruction are a more natural fit for online instruction.

How to Implement Online Instruction

Another area of research inquiry concerns how practitioners can most-effectively implement specific strategies of online instruction in order to achieve instructional goals. An example is Volery and Lord's identification of three critical success factors in the implementation of online instruction: "technology, the instructor, and the previous use of the technology from a student's perspective" (2000, p. 216). Other researchers, such as Morrow

and Shi (2006), have made more specific recommendations for when and how to implement certain online learning tools. Still other studies have narrowed the scopes of their inquiries in order to make detailed recommendations about the implementation of specific tools in order to obtain pre-defined results. To again use the topic of asynchronous discussions as an example, Ngwenya, Annand, & Wang (2004) plus Vonderwell (2003) recommend specific wording in the questions posted to online forums in order to improve student engagement, Sendag and Odabasi (2009) recommend the use of problem-based learning discussions to boost critical thinking, and Li et al (2010) recommend implementing guaranteed response times in order to boost student satisfaction.

Some of the most interesting future research in this vein is likely to come from MOOC publishers, since have access to so much data about how students engage with their learning environments. An example is Guo, Kim, and Rubin's (2014) analysis of how students respond to online videos, which shows that on average students prefer short (<6 minute), informal videos which show both talking heads and tablet drawing. On the whole, these studies which explore best practices for certain types of online instruction offer useful guidance in terms of implementing specific online learning strategies, but still do not offer much in the way of comparison between strategies. For a school that is already locked in to a specific strategy for online instruction, the recommendations that these studies contain may be extremely valuable, but they do not provide as much guidance about which strategy to adopt in the first place as could be desired.

Outcomes for Different Student Groups in Online Instruction

On the other hand, if a school or course serves a fairly homogenous student population, its instructors may benefit from the knowledge presented by the growing number of studies which investigate how different student groups fair in online educational settings. In general, these studies seek to compare the outcomes for certain student groups in online classes relative to their peers in traditional learning environments. Xu and Jaggars (2013b), for example, have shown evidence that low-GPA, Male, Black, and younger students may all perform worse in online versus traditional classes. Along the same lines, Figlio, Rush, and Yin (2010) also showed evidence of underperformance among low-GPA, Male, and Hispanic students taking online classes. In terms of learning styles, research has shown that student satisfaction and perceived learning outcomes within online discussions are highest for students with visual and read/write learning styles (Eom, Wen, & Ashill, 2006). Unsurprisingly given this result, other research has indicated that visual and read/write learners tend to self-select into online courses (Drago & Wagner, 2004). Another study of note is Offir, Lev, and Bezalel's (2008) work which shows that low-cognitive-ability students are less successful in asynchronous versus synchronous online classes. This last study is significant because it offers a direct comparison between different online instructional techniques—which is the subject of the last group of studies in this review of the literature.

Comparisons of Online Strategies for Instruction

The final broad category of studies on online learning deals with comparisons between different strategies of online instruction. It is this last category of study which is most pertinent to this paper given the nature of its inquiry. Unfortunately, this small group of studies, when

taken on the whole, present something of a mixed picture with often contradictory or inconclusive results. There are however, a number of studies in this category which offer direct comparisons between online instructional strategies and are able to identify (and replicate) a relatively clear superior strategy in terms of the outcomes or benefits measured. Examples include studies which have compared hybrid (asynchronous and in-person) instruction to wholly asynchronous online instruction and have demonstrated superior student learning for hybrid instruction (Ge, 2012) as well as higher student satisfaction (Ausburn, 2004; Rovai & Jordan, 2004).

Other notable studies in this vein include Choi and Johnson's (2005) comparison of video versus text-based online instruction, which showed superior motivation towards attentiveness and memorization for video-based lessons. Meanwhile, both Johnson (2006) and McDonald (2002) cite numerous advantages of asynchronous instruction over synchronous instruction. On the hand, still other studies comparing asynchronous and synchronous methods of instruction have shown no significant difference in terms of student learning outcomes (Cleveland-Innes & Ally, 2013; Skylar, 2009). There are also several examples of researchers who have compared asynchronous versus synchronous online instruction and have found that there are positives and negatives to each approach, often depending on context. These include Hrastinski (2008a), who found student communication in asynchronous online courses to be more thoughtful and complex while their communication in synchronous courses tends to be more social and interactive. Oztok et al (2013) also found similar results to Hrastinski in their analysis of asynchronous and synchronous online communication. Leaving aside instructional effectiveness, Laws, Howell, & Lindsay (2003) point out that asynchronous

instruction is capable of scaling more effectively while synchronous instruction, though less expensive to implement, is highly faculty-intensive.

In short, the existing comparisons between online instructional strategies offers a myriad of different perspectives, and the amount of literature is still too small to get a good sense for most trends in the results. Additionally, most of the comparisons are made between just two online instructional strategies (i.e. hybrid vs. asynchronous or videoconferencing vs. text-based discussion forums) and it is therefore challenging to gain insight into which of the many strategies may be most effective overall, in spite of the usefulness of these one-to-one comparisons. Among the key takeaways are some reasonably strong indications that hybrid instruction performs strongly versus other online strategies—and compared to traditional instruction as well (Means, Toyama, Murphy, Bakia, & Jones, 2010; Rovai & Jordan, 2004). Nevertheless, there is hardly enough evidence at this time to declare hybrid the "best" of online instructional strategies, and there is still a great deal to be learned about which online education strategies may be most effective in terms of their effects upon many important student outcomes. This study builds upon the work of the scholars in this last category of research by performing a simultaneous, side-by-side comparison of several major online instructional strategies and by evaluating differences in students' learning, satisfaction, participation, and engagement across them all.

Chapter Summary

Among the main takeaways from this chapter should be the understanding that there is still a great deal to learn about online education. Unlike traditional classroom instruction, which has existed for decades in a format that is at least recognizable across generations

(Cuban L. , 1993), online education is not only relatively new, but constantly evolving. In similar fashion to genetic evolution, online education could be described as being in a phase of "radiation" where a diversity of forms are currently proliferating. This presents challenges for practitioners and researchers alike for, as Kearsley (1997) points outs:

> Because there are so many programs available now for CMC [computer mediated communication]; it is difficult to evaluate them and decide which one(s) to use. Of course, over time personal experience and research studies will provide guidance but at present the technology is evolving too quickly for there to be much of either. (pp. 6-7)

Unfortunately, the rapid pace of adoption for online instruction (Allen & Seaman, 2014) is unlikely to slow enough in the near term to allow the research to catch up. Just as confounding is the difficulty in predicting which new instructional approaches and technology will be in use even just ten years from today. For these reasons, it is particularly important that research begins to form an understanding of the comparative effectiveness of some of the most fundamentally different strategies of online education, such as asynchronous versus synchronous versus hybrid instruction. While software platforms will undoubtedly change, issues like whether participants are engaged at the same moment in time and whether or not they ever meet in person are likely to remain as fundamental distinctions.

Perhaps the most important finding from the existing research is the over-arching theme that online education, generally-defined, appears to be at very least a viable alternative to traditional instruction. In the emerging field of online education research, this was the obvious first thing to investigate and the large meta-analyses conducted by Means et al (2010), Bernard et al (2004), and Jahng, Krug, and Zhang (2007), among others, offer ample evidence to

this effect. Their work, along with Russell (1999), has established the baseline from which further research can proceed to identify the most effective online instructional strategies in terms of improving student outcomes, in order to better inform those educators who have elected to develop online courses.

Chapter 3: Methods

Introduction

Organization of the Chapter

This chapter details the methodology used in this study, beginning with a brief, general overview of the study in order to orient the reader in subsequent sections. This is followed by a discussion of the experimental research design and the reasoning behind it given the study's objectives. The next section explains the key independent variables, or factors, that were included; more specifically, the online instructional strategies that were used to deliver the lesson to each treatment group. After this is a section which details the dependent variables measured for each group and the instruments used to collect them. Next comes a description of the study's participants followed by a walkthrough of the data collection procedures. The chapter ends with a discussion of how the resulting data were analyzed.

Restatement of the Study's Purpose, Significance, and Research Question

Strategies for online education can take a variety of forms with several significant distinctions in terms of these strategies' formats and approaches. Although there is ample evidence which suggests that online education, broadly-defined, can be as effective as traditional, classroom instruction, there has so far been little investigation into how distinct strategies for online instruction may differ in their effects upon student outcomes. The purpose of this study is to measure and compare the effectiveness of various strategies for online instruction in terms of their abilities to affect several key student outcomes. The ultimate goal of the study has been to make progress towards answering the following research

question: "Are there differences across students' comprehension, satisfaction, participation, and engagement according to whether a lesson is taught in class, in a hybrid format, online using videoconferencing, online using pre-recorded video, or online using text and images?" By addressing this issue and providing some evidence about whether certain online instructional strategies may be more effective than others, the hope is that this study will allow educators to be able to make better-informed decisions when designing and building online educational content.

Introductory Summary of the Methodology

The following paragraph provides a brief summary of the research methods which will be discussed at greater length in the subsequent sections of this chapter. This section is intended to give context to the more detailed information to follow.

In short, this study has used an experimental design in order to compare the effectiveness of several different strategies for online instruction in terms of students' comprehension, engagement, participation, and satisfaction in a single, hour-long lesson. The experiment began with the recruitment of undergraduate student participants from one U.S. university campus to take part in a stand-alone lesson that was not affiliated with an existing class. After collecting demographic information on the participants through an online survey, each participant was randomly assigned to one of five treatment groups. Then, across all five groups, the same lesson (allowing for slight differences in delivery) was conducted by the same instructor concerning a subject about which the students had little prior knowledge. While one control group was taught in a traditional classroom format, the four other groups received instruction delivered via different online instructional strategies which represent the current

major paradigms of online instruction in present-day U.S. higher education. Each group was then evaluated separately across several key student outcomes using similar assessment procedures for all groups. Finally, the results of the evaluations were compared across all five groups in order to determine whether there were significant differences in student outcomes depending on the instructional strategy employed.

Pilot Study

A study similar to the one outlined throughout this chapter was carried out at the same institution in April 2013 using fewer variables and a considerably-smaller group of participants, but most of the same methodology. The results of this earlier experiment served to validate the effectiveness of this study's experimental design, the lesson used, and the instruments employed to assess comprehension and measure student satisfaction. In the pilot study, seventy-two participants were divided into three groups, each of which received instruction using one of three strategies: face-to-face, synchronous videoconferencing, and asynchronous video. The results of the pilot study showed no significant differences in terms of comprehension and satisfaction across instructional strategies, but the small number of participants resulted in the pilot experiment having low statistical power. Nevertheless, this earlier experiment was a useful practice run for the procedures that were used in this study—demonstrating both the effectiveness of the measurement instruments and the ability of the instructor to deliver a nearly identical lesson in different formats.

Experimental Research Design

Discussion of Experimental Design

Keppel (1991, p. 5) defines experimentation in the following way: "Differential treatments are administered to different groups of subjects … and performance on some response measure is observed and recorded following the administration of the treatments." In a true experiment, subjects are randomly assigned to different treatment groups. As Kirk (1982, p. 24) explains, "Through random assignment, a researcher creates two or more groups of participants that at the time of assignment are probabilistically similar on the average." Naturally, there are bound to have been small differences between the subjects in each of these randomly-assigned groups, but the differences are very likely to have been small enough that they did not significantly affect the results of the experiment (Wallen & Fraenkel, 2000). This randomized assignment of subjects to treatment groups is, in fact, an important attribute of experimental design since it allows the researcher(s) to control for differences in subjects and to avoid confounding factors such as selection bias (Keppel, 1991). Because of the advantages offered by experimental design in educational research, Campbell and Stanley (1966, p. 2) even go so far as to refer to experimentation "as the only means for settling disputes regarding educational practice, as the only way of verifying educational improvements, and as the only way of establishing a cumulative tradition in which improvements can be introduced without the danger of a faddish discard of old wisdom in favor of inferior novelties."

Application of Experimental Design to the Study

In following a between-subject experimental design (Keppel, 1991), participants in this study were randomly assigned to one of five treatment groups. Due to an attrition rate of 22%

in the pilot study, it was deemed inadvisable to conduct participant matching, since a similarly-high attrition rate might have led to undesirable differences across the treatment groups if not randomly-assigned (Creswell, 2009). Because each treatment group differed primarily in just the format of instruction that they received, this study can be classified as a single-factor factorial design (Glass & Stanley, 1970). In addition, since the instructional formats between groups represent a single, key categorical variable, but several different outcomes were measured, the experiment is multivariate for only the dependent variables.

Although some non-treatment independent variables concerning the participants' past exposure to the subject matter and to online instruction were recorded, this information does not constitute a pre-test for the measured outcomes. Campbell and Stanley (1966, p. 25) note that a pre-test is not necessary in a true experiment, stating, "The most adequate all-purpose assurance of lack of initial biases between groups is randomization. Within the limits of confidence stated by the tests of significance, randomization can suffice without the pre-test." Hence, this study follows a post-test-only control group design (Campbell & Stanley, 1966) which, as Creswell (2009) points out, eliminates the possibility of the pre-test confounding the results of the dependent variables. Following Campbell and Stanley's (1966) approach, a visual model of this study's experimental design is presented below:

Group 1 R——————X_1——————O

Group 2 R——————X_2——————O

Group 3 R——————X_3——————O

Group 4 R——————X_4——————O

Group 5 R——————X_5——————O

In this model, meant to be read left to right as a time-series, the participants are all randomly assigned (designated by R) to one of five treatment groups (1-5). Each group is exposed to one of five instructional strategies, $X_1 – X_5$, as treatments, and then all outcomes are measured together.

Online Instructional Strategies as Treatments

Group C: Classroom Control (Face-to-Face)

This group attended a traditional, in-person class for 45 minutes of lecture-based instruction. In this session, the students were permitted to ask questions and make comments during and after the lecture—as is the case in most typical classroom environments. In total, participants in this section made fifteen comments during the course of the lecture. The lecture was followed by 15 minutes of in-person discussion in which the instructor asked the students to provide their opinions on several issues related to the content of the lecture.

Group C served as the control group for the experiment. While the intent of this study is primarily to compare different online instructional strategies, the results from Group C were used to benchmark the results from the other groups against the traditional, classroom format for instruction.

Group S: Synchronous Videoconference

Students in Group S logged in to a videoconferencing platform at a pre-designated time from whatever location they chose. They watched the same instructor deliver a nearly-identical 45-minute lecture as was delivered to the in-person group (allowing for slight differences in spoken content, interruptions, and pace) live online via the Blackboard

Collaborate online meeting tool. During the lecture, students were able to ask the instructor questions and conduct discussions amongst themselves via live chat. In total, participants in this section made twelve comments during the course of the lecture, though not all were addressed individually by the instructor. The live online lecture was followed by 15 minutes of live discussion led by the instructor using the videoconferencing system's interactive tools which covering the same questions and issues presented to Group C.

Participants assigned to Group S experienced online instruction in a similar fashion to students whose institutions make use of web meeting tools[1] for live online instruction. While not as common as asynchronous online instruction, a majority of higher education institutions with distance learning programs do use synchronous web-conferencing tools to some extent (Parsad & Lewis, 2008; Burnett, 2003). Furthermore, there is disagreement in the existing literature concerning the effectiveness of synchronous online instruction, with some studies such as Bernard (2004) characterizing it as the least-effective of major online instructional strategies, while others, such as Bullen (1998), suggest that it may be the most effective. Also of note is the idea suggested by some education researchers that synchronous instruction may be the online strategy which most closely aligns with traditional best practices for classroom instruction (Lieblein, 2000; Morrow & Shi, 2006).

[1] Such as WebEx, Wimba, Adobe Connect, Blackboard Collaborate, Polycom, Cisco TelePresence, Skype, and others

Group H: Hybrid

Hybrid, or blended, learning typically involves a combination of asynchronous content, such as pre-recorded video lectures, coupled with in-person classroom discussions (Bishop & Verleger, 2013). Students in Group H were given access to an approximately equivalent 40-minute lecture as Group C (again, allowing for slight differences in delivery), but in the form of a pre-recorded video. This video lecture was 5-minutes shorter than the lectures for Groups C and S primarily due to a lack of interruptions and questions. Participants in Group H were able to watch this lecture whenever they choose to within the course of one week's time. They then attended an in-person class for 20 minutes of discussion with the instructor which dealt with the same issues and questions presented to the other groups.

One compelling reason to include a hybrid treatment group in this study is that, in quite a few other studies, hybrid instruction has outperformed both fully online and traditional instruction (Means, Toyama, Murphy, Bakia, & Jones, 2010; Rovai & Jordan, 2004; Bishop & Verleger, 2013; Ge, 2012). While this does not mean that other online strategies have no other advantages over hybrid courses, it is useful to evaluate the question of whether hybrid instruction can outperform other online strategies in the transmission of a single, identical lesson. More than just using a blend of online and face-to-face instruction, the treatment given to Group H generally followed a flipped classroom format in which the lecture was be delivered outside the classroom while the time in-person was devoted to discussion (Bishop & Verleger, 2013). The flipped classroom is one of the most popular emerging trends within higher education instruction (Fitzpatrick, 2012) and it is therefore useful to compare its outcomes with other widely-used strategies of online instruction.

Group V: Asynchronous Video and Forums

Another major strategy for online education involves the use of pre-recorded videos for lecture delivery coupled with asynchronous online discussion boards for student and instructor interaction. Group V was given access to the exact same 40-minute pre-recorded lecture as Group H to watch sometime within a one-week period. These students were then asked to participate in an instructor-moderated online discussion forum in which the same issues and questions were presented by the instructor in an asynchronous format.

It is important to include a treatment for asynchronous instruction because this format is one of the most common online instructional strategies. Asynchronous instruction is facilitated by all major Learning Management Systems[2] and also is the strategy of choice for most MOOC providers (Skiba, 2012). Furthermore, as Hrastinski (2008b) points out, there are some students who choose to study online specifically because they can do so asynchronously—which allows them flexibility in their schedules. With regards to the use of pre-recorded video, this is among the most-popular mediums for the delivery of instruction outside of the classroom (Parsad & Lewis, 2008). Laurillard (2013, p. 103) refers to video as having the ability "to bring together experience and description of that experience and, being self-paced, can enhance further with the opportunity for students to reflect on what they are doing."

[2] Blackboard, Moodle, Canvas, Sakai, eCollege, Desire2Learn, and many others

Group T: Asynchronous Text and Images with Forums

Group T received the same fundamental lecture information as the other groups by reading online text and viewing slideshows housed in the Moodle Learning Management System. The students in Group T were then asked to participate in an online discussion forum with the same design provided to Group V. In effect, the instruction provided to Group T was equivalent to that of Group V, but with the media used to deliver the lecture changed from video to text-based documents.

The reasoning behind the inclusion of this treatment was to evaluate the effect which the media used for asynchronous instruction may have upon student outcomes. Differences in outcome measures for different media, or the lack thereof, may offer evidence to refute or support Clark's (1983) assertion that the type of media used does not affect the quality of instruction. Another reason to include text-based instruction as a treatment is that the delivery of asynchronous online lessons primarily through documents and non-narrated slides remains common practice in online higher education (Parsad & Lewis, 2008). While perhaps not as visually-captivating as a well-produced video, text-based online instruction has served as something of a link for distance education programs that are transitioning to more modern online instructional strategies (Taylor, 2001). It is also worth noting that text-based asynchronous instruction is typically the least-expensive and least-difficult online instructional strategy to implement (Rumble, 2001), so it is understandable that it remains commonly-used.

Cross-Group Similarities and Differences

Across the five treatment groups, all of the lessons were delivered by the same instructor—a faculty member in the School of Management at the same school from which the

participants were recruited. This instructor was the same faculty member who taught the lessons for the pilot study. Although he had some experience with being video recorded while teaching and had used videoconferencing to conduct online office hours, he had not taught an online or hybrid course at the time that this study was conducted. This instructor was selected because of his expertise and experience teaching a subject —international intellectual property law—which is relatively easily explained during the course of a single, one-hour lesson, but about which most undergraduates students are unlikely to have much significant exposure. A lesson on a subject of law was appropriate because other studies and articles have indicated that legal studies can be taught effectively online and do not show indications of significantly different results from other academic subjects when taught in an online environment (Shelley, Swartz, & Cole, 2007; Marcel, 2002; Miller, 2004). This lesson was delivered to all five treatment groups using the same PowerPoint slide deck with as little deviation in the delivery of the lecture as could reasonably be expected.

Aside from the instructional strategy employed, the biggest differences in the delivery of the lectures between groups were due to the questions asked and comments made by participants in Groups C and S. While none of these class contributions took the lecture in a significantly different direction, in both cases they did extend the length of the lecture by about five minutes beyond what was experienced by Groups H and V. It is also possible that the instructor moved somewhat faster through some slides in order to make up for time spent answering questions. The only other notable difference between the groups was that the students in Group S twice lost their video image of the instructor when his screen saver

activated. These disruptions were minor, however, and not atypical of what might be expected in any class conducted via videoconference.

In terms of execution, Groups C and S (in-person and synchronous) had their lessons on the same day, while Groups V and T (asynchronous video and text) were given a week-long window in which to view the lecture and participate in the asynchronous discussion. Group H (hybrid) also had one week to view the lecture content and then participated in the classroom discussion portion of their lesson at the end of the week—on the same day as Groups C and S. This timing and format were intended to mirror the realities of how synchronous and asynchronous courses are taught online. In all groups, participation in the discussions was encouraged, but was not described as mandatory. If members of the groups with asynchronous content chose to access the content multiple times or browse the internet during viewings, this is not a cause for concern because it is an accurate representation of possible student activities in a real asynchronous learning environment.

Outcome Measures

Assessment of Comprehension

In order to measure their comprehension of the material presented in the lesson, participants in all treatment groups were asked to take a short quiz during the week following their lessons. This quiz consisted of fifteen multiple-choice or short-answer questions that measured memory and comprehension of the content of the lesson. The quiz questions were developed by the instructor who taught the lessons, and were drawn from questions that he typically assigns his students after delivering a similar lecture in the classroom for a credit-

bearing course that he teaches. The questions and answers for this assessment are listed in Appendix B.

Numerous researchers have pointed out the important role that frequent assessment can play in online learning (Kirkwood & Price, 2008; Tshibalo, 2007; Gikandi, Morrow, & Davis, 2011). For some online learning strategies, regular assessments can provide a predictable "check-in" that substitutes for class time and can help to keep students on track in learning environments in which it might otherwise be easier to fall behind (Gaytan & McEwen, 2007). In addition, some research suggests that by emphasizing the important points from online lessons, assessments can improve learning outcomes for online students (Angus & Watson, 2009). Making use of an online assessment in this study, therefore, serves the dual purpose of measuring participants' comprehension of the lecture content and of employing a frequently-used component of many online education strategies.

Survey of Satisfaction

At the same time that they completed the assessment, participants in this study were also asked to fill out a short online survey to measure their satisfaction with the lesson. This survey consisted of fifteen Likert scale questions with levels ranging from one through five. It can be found in Appendix C. Within the survey were five categories of three questions each. These categories consisted of questions pertaining to participants' satisfaction with the subject matter of the lesson, the instructor, the lesson format, their interactions during the lesson and subsequent discussion, and the overall learning experience.

It was important to measure student satisfaction across the treatment groups because there are considerable differences of opinion about whether students in online courses may be

more or less satisfied about their learning experiences. Some researchers assert that satisfaction rates among online students are higher (Eom, Wen, & Ashill, 2006), while others cite evidence which shows that they are lower (Benoit, Benoit, Milyo, & Hansen, 2006; Cole, Shelley, & Swartz, 2012), or are roughly the same (Allen, Bourhis, Burrell, & Mabry, 2002) when compared to traditional classroom instruction. This study seeks to understand whether these differences in measured rates of satisfaction among online learners may be the result of different strategies of online instruction.

Student evaluation surveys are a standard form of measuring the effectiveness of instruction, and their validity has been documented and widely supported by research (Centra, 1993). More recently, online course evaluations have gained in popularity (Hoffman, 2003) and there is ample evidence that the use of online surveys for course and instructor evaluations are effective and do not produce different results when compared to paper-based surveys of the same class (Dommeyer, Baum, W., & Chapman, 2004; Couper & Bosnjak, 2010; Hardy, 2003).

As Creswell (2009) points out, surveys provide an efficient means of gathering perspectives from a sample which can then be generalized to a population. Per Fowler (2013), surveys also provide standardized measurement, meaning that information collected from subjects can easily be quantified and compared. In terms of survey design, Fowler (2002) and de Vaus (1991) both recommend closed questions to provide more concrete and quantifiable information. Likert Scale questions, in particular, are a “reliable and sensitive mirror of attitudes” (Moser, 1958, p. 239). One common threat to validity in many survey designs, nonresponse error (Fowler, Survey Research Methods, Fifth Edition, 2013), did not influence the results of this study, since completion of the survey was the final and relatively less-arduous

task required for participation and all participants who engaged with the lesson also filled out the survey.

Quality and Frequency of Engagement Measured by Discussion Responses

As noted in the section describing the instructional strategies which served as treatments, all groups were prompted with the same discussion questions following their lessons. The in-person, hybrid, and online synchronous groups had the questions posed to them live and conducted their discussions in real time, while the two asynchronous groups carried out their discussions in an online forum. To evaluate the quality of discussion responses, the responses and comments from all treatment groups were transcribed from recordings of the live sessions or copied directly from the online forums. Comments made and questions asked during the lectures of Groups C and S were included in this transcription. After removing identification from the transcribed/copied responses, but including context, an experienced teaching assistant from another institution who is familiar with the subject matter used a rubric adapted from the Eberly Center for Teaching Excellence at Carnegie Mellon University (Chang) to grade the quality and insight displayed by each of the responses on a scale from one to four (see Appendix D). The total number of comments from each group, as well as the number of students who participated in each group's discussion were also recorded.

A number of experts in online education have cited the important role that discussion forums play in many online learning environments (Roblyer & Wiencke, 2003; Harman & Koohang, 2005). In addition, Bean and Peterson (1998) argue for the use of rubrics as an effective means of evaluating online discussion contributions and as a way for instructors to help encourage meaningful class participation. By using a third party to grade the quality of

discussion responses without knowing their origin, this study has sought to provide a better comparison of the level and quality of engagement across instructional strategies.

This comparison may prove to be significant for practitioners since communication and interaction in online courses represents one of the most important departures from traditional classroom teaching. Instructors who are accustomed to a great deal of back-and-forth dialogue with their students in class may naturally feel some apprehension or skepticism about the quality of communication in an online environment. When selecting between online strategies, these instructors may gravitate towards synchronous tools for communication as it more closely resembles dialogue in person (Lieblein, 2000). At the same time, some researchers have pointed out distinct advantages of asynchronous communication, including the opportunity to allow for full-class participation without time constraints (Li, Finley, Pitts, & Guo, 2010; Vonderwell, Liang, & Alderman, 2003). Others have noted that students tend to put more thought into written, asynchronous responses—resulting in higher-quality, if less spontaneous, responses (Hrastinski, 2008a; Johnson G. M., 2006).

Participation/Attendance

According to data collected by Allen and Seaman (2014), 41% of chief academic officers in U.S. higher education institutions believe that retention is a bigger problem in online classes than in traditional classroom environments. This concern is borne out in other research, much of which points to higher drop rates in online classes (Gleason, 2004; Diaz, 2002; Brady, 2001). While this study will not be able to offer much evidence on this issue through the examination of just a single lesson, it is worthwhile to examine whether there are differences in attrition rates across the treatment groups from registration through the conclusion of the study. While

this study cannot hope to simulate the circumstances which affect participation and attrition during an entire course, evidence of significant differences in attrition might suggest that higher levels of accountability and/or convenience inherent in the different online instructional strategies may influence students' likelihood of participation.

Population and Sample

The participants in this experiment were 303 undergraduate students enrolled at a major U.S. West Coast public research university. This institution is located in a metropolitan area and is highly selective in admissions. In terms of demographic, Asians and Pacific Islanders make up the largest racial group followed by Caucasians and Hispanics. Approximately one eighth of undergraduates at this institution are from outside the U.S. The choice to conduct this study at a large, public university is appropriate because demand for online courses in the U.S. is highest at public universities (Allen & Seaman, 2009) and because leaders of public universities more often believe that online education is critical to their long-term strategies (Allen & Seaman, 2011). In addition, public doctoral research universities in the U.S. are currently at the forefront of the MOOC movement and offer the most of these large online courses (Allen & Seaman, 2013). In spite of these trends, the university at which this study was conducted has not been heavily involved in online instruction and offers few online classes when compared to many similar institutions of its size.

Aside from their mutual enrollment as undergraduate students at the same university, the students who were recruited for this lesson did not necessarily share any other common characteristics. This study was conducted as a single, stand-alone lesson and was not affiliated with any existing class or activity. Students who registered for the study were recruited

through research Listservs, online forums affiliated with the university, the campus newspaper, and with flyers posted around the university. All of these recruitment materials were linked to an online survey through which the subjects could indicate their eligibility for and interest in becoming participants in the study. This online survey was shut down once 425 eligible students had registered to participate.

In addition to certifying their enrollment as bachelor's degree candidates at the university in question, another condition for participating in the study was that subjects were required to certify their availability for the day and times that all live sessions were to take place. Participants could not also have been subjects in the pilot study, since the same lectures and many of the same instruments for collecting data were reused. In addition, each participating student received $30 as compensation for their time. In total, 425 students registered to take part in the study, of whom 303 actually participated.

In short, although the participants were randomly assigned to treatment groups, this study did not engage in random sampling for the recruitment of participants, and there may be legitimate concerns about whether the participants who signed up for this study are even fully representative of the student body at their institution. Further analysis of the demographics of the study participants and the university population from which they were drawn is presented in Chapter 4. In spite of the potential lack of representativeness of the overall U.S. undergraduate population from the study participants, the homogeneity provided by drawing the sample from a single institution provides the benefit of reducing within-factor variability across treatment groups, allowing for more-accurate insight into the effects of the different instructional strategies upon the measured outcomes (Glass & Stanley, 1970).

Chronology of Data Collection Procedures

Upon their recruitment, each subject received a short online survey (Appendix A) to gather demographic information. At this time, they were also asked to verify their availability for all possible live activities associated with any treatment group. This required that all students be free for the greater part of a Friday afternoon in the middle of the fall quarter, since all live lecture and discussion activities for Groups C, S, and H were to take place on the same day, but at different times, in order to keep the groups separated. Following their completion of this survey and their random assignment to treatment groups, each student received an email detailing instructions for their specific treatment exactly one week before the live sessions were to be held. With this email, participants in the hybrid and asynchronous groups (H, V, and T) also received immediate access to their online instructional content, which they were able to view at any time during that week. The two asynchronous online groups (V and T) were also able to participate in their respective discussion forums throughout the week.

The live components of the treatments for Groups C, S, and H all took place on Friday afternoon at the end of this week. First, from 3:00 to 4:00pm, Group S conducted their online lecture and discussion remotely via videoconference. Then, at 4:30, Group H participated in a 20-minute live discussion in an auditorium-style lecture hall located on campus. Finally, Group C completed their lecture and discussion from 5:00 to 6:00pm in the same lecture hall used by Group H. All of these live activities were video recorded for later analysis. As soon as all live instructional and discussion activities were concluded, Groups H, V, and T lost access to their online lecture materials and participants in all groups received access to the same online quiz and questionnaire. Though they were given a full week to complete the questionnaire and quiz,

there was a time limit placed on the quiz so that, once any participant opened it, they would need to submit their answers within thirty minutes of the time when it was opened. At the conclusion of the second week, the assessment and questionnaire were both closed and the results analyzed.

Data Analysis

Independent Variables

The key independent variable in this study is the strategy of instruction employed, which was coded as dichotomous variables for analysis, with Group C (face-to-face) used as the reference group for regressions. In addition, nine demographic variables from the pre-study questionnaire (Appendix A) were also included in the regression analyses. These variables, which are also categorical in nature, and are listed below in Table 3.1. Note that, due to low response rates for some questions, certain responses from the pre-study questionnaire were combined into a single category for analysis. This step was also necessary in order to reduce the number of predictor variables in the regression models – particular in the models used to test for interaction effects. For instance, responses indicating GPAs below 2.50 and from 2.50 to 2.99 were combined into a single category for any GPA below 3.00. In addition, the responses for Undergraduate Major were combined in order to distinguish between STEM and non-STEM majors in order to provide better comparisons to the literature (Xu & Jaggars, 2013b).

Table 3.1. Demographic Independent Variables

Categorical Variable	Reference Group	Comparison Group 1	Comparison Group 2
Age	<18	18-19	>19
Gender	Male	Female	
Race	Caucasian	Asian	URM
1st Language	English	Other	
Undergraduate Major	non-STEM	STEM	
Undergraduate GPA	<3.00	3.00-3.49	>3.49
Experience in Online Classes	None	Some	
Experience with Intellectual Property Law	None	Some	
Interest in the Subject Matter	Slight or None	Moderate or Greater	

Age, gender, race, GPA, and students' prior experience with instructional technology have all been cited as significant factor in the success of students taking online courses (Caspi, Chajut, & Saporta, 2008; Figlio, Rush, & Yin, 2010; Volery & Lord, 2000; Xu & Jaggars, 2013b). Major, whether English was the first language, and experience with and interest in the subject matter were all included as factors to control for since they could conceivably alter the outcome measures of the study.

Dependent Variables

Ten dependent variables were examined to look for differences across the treatment groups. These dependent variables are listed below in Table 3.2. Note that the averages from five categories (3 questions each) of the post-lesson satisfaction survey (Appendix C), in addition to the overall survey average, were each analyzed as a separate dependent variable.

Table 3.2. Key Dependent Variables

Dependent Variable Abbreviation	Range	Description
Participation	0-1	Measure of whether a registrant completed all elements of the study
Post-Lesson Assessment of Comprehension (quiz score)	0-100	15 questions, 6.67 points each (see Appendix B)
Commenting (discussion participation)	0-1	Measure of whether a participant made one or more comments during the lesson or discussion
Average Score per Comment	1-4	Average score for group comments (see Appendix D for the rubric used to score the comments)
Satisfaction with Subject Matter	1-5	Average of scores on questions 1-3 from the post-lesson survey of satisfaction (see Appendix C)
Satisfaction with Instructor	1-5	Average of scores on questions 4-6 from the post-lesson survey of satisfaction (see Appendix C)
Satisfaction with Lesson Format	1-5	Average of scores on questions 7-9 from the post-lesson survey of satisfaction (see Appendix C)
Satisfaction with Interactions	1-5	Average of scores on questions 10-12 from the post-lesson survey of satisfaction (see Appendix C)
Satisfaction with Learning Experience	1-5	Average of scores on questions 13-15 from the post-lesson survey of satisfaction (see Appendix C)
Overall Survey Average	1-5	Average of scores on all questions (1-15) from the post-lesson survey of satisfaction (see Appendix C)

Statistical Techniques

The data analysis for this study began with gathering descriptive statistics for both the demographic and outcome variables. The means and standard deviations across treatment group categories were then used to calculate effect sizes and statistical power to inform the later analysis of each of the different dependent variables. Next, ANOVAs were used to look for evidence of the existence of differences in the numeric measured outcomes across the treatment groups. For the dichotomous dependent variables of study completion and participation rates, chi-square tests were used instead of ANOVA in order to look for differences between treatments. For these ANOVAs and chi-square tests, the Bonferroni correction was used to counteract the problem of multiple comparisons. Multiple linear regressions were then employed to analyze each of the eight dependent variables with numeric outcome measures, while multiple logistic regressions was used to analyze the two dichotomous outcomes.

For eight of the ten dependent variables, these regressions included both the treatment groups and demographic variables. Although this study is primarily concerned with differences in student outcomes based on the strategy of instruction employed, demographic variables were included to investigate the possible effects of these characteristics upon the dependent variables and to inform later investigations into the possibility of interaction effects. Variables from the pre-study questionnaire were not included in the analysis of discussion participation or average comment score because it was not always possible to ascertain the identity of each speaker from the recordings of the live sessions from Groups C, S, and H. Regardless of the possible effects of the demographic variables on the regressions, the randomized assignment of

participants to the treatment groups should have controlled for most between-group differences (Wallen & Fraenkel, 2000).

As a final step in the analysis of the data, each of the outcome measures (with the exception of comment rates and scores) were tested for the presence of interaction effects between the treatment group variables and the demographic variables. This was accomplished by running seventy-two separate multiple regressions—i.e., a separate test for each of the nine demographic variables on each of eight outcome measures for which demographic information was available. All of these regressions included as independent variables the categories for each treatment group, the categories for all demographic variables, and the interaction terms obtained by multiplying the treatment group variables by a specific demographic variable category. While the performance of so many regressions may have exposed the results to Type I errors, it was deemed that there was more to be gained by investigating all possible interactions while recognizing the possibility of Type I errors. This issue is discussed in greater detail in the section on interactions in Chapter 4.

Threats to Validity

One issue which could throw some doubt on the validity of the results is the possibility that some students cheated on their assessments or conducted further research into the subject matter after the conclusion of their lessons but before taking their assessments. While this is certainly a possibility, the risk of this event was equal across groups. In other words, the randomized assignment of participants to treatment groups should have as effectively controlled for the possibility of academic dishonesty as it did for pre-exposure to either the subject matter or the method of instruction used in the lesson. Since all groups had access to

the assessment in the same format, none of the treatment methods should have resulted in a greater propensity to cheat.

Another potential concern has to do with the timing of events in the experiment. While the fact that all activities in the experiment occurred during the same two-week span should address potential concerns about campus events having an influence on some treatment groups more than others, it could still be argued that the greater time granted to the asynchronous groups to view their lessons and participate in discussions might skew the results. In fact, the extra time allotted to the asynchronous groups is actually a better representation of the true differences between asynchronous instruction and other strategies in a real-life instructional environment. This point can likewise be used to address concerns about participants in the asynchronous groups' ability to review their lessons, since this would also be true in any online asynchronous course.

It could also be pointed out that the design of the experiment makes it impossible to distinguish between the methods employed in conducting the lecture or the discussion in terms of their respective effects upon the measured outcomes. If, for instance, the hybrid group outperforms the videoconferenced group, this might be either because pre-recorded video is superior to synchronous video or because online discussion forums provide a better way to reinforce concepts from the lecture than the online chats which occur during a videoconference. This point is certainly valid in that this study does not offer ways to distinguish between the effects of the lecture versus the discussion. That noted, the combinations of lecture and discussion formats were chosen based upon the most frequently-used combinations in major strategies for online instruction. In face-to-face and

videoconferenced lessons (Groups C and S), the discussion most naturally progresses from the lecture in the same format. In hybrid classes (Group H), the lecture is typically delivered asynchronously online then followed by an in-person discussion (Bishop & Verleger, 2013). And, in most asynchronous online courses (Groups V and T), the discussion forum is the primary means of communication between the students and the instructor (Roblyer & Wiencke, 2003; Harman & Koohang, 2005). Thus, it should be reaffirmed that this study seeks to compare several major strategies of online instruction, each of which are made up of dual components, as opposed to comparing those components themselves.

Other potential threats to validity include the possibility of the instructor influencing the results of the experiment. While this could be a legitimate concern, the instructor who was selected to deliver the lesson for this study has little personal stake in any online learning strategy. Though he has some experience with synchronous instruction over videoconference, he is primarily a classroom instructor. In addition, since his performance was recorded, it was possible to analyze the similarities between his deliveries across treatment groups and overall he showed remarkable consistency in delivering an almost identical lecture across the different instructional strategies. Likewise, the teaching assistant who graded the discussion responses had even less reason or opportunity to try to influence the results of the experiment since he is unaffiliated with the research or the school. Also, this teaching assistant was provided with only a basic summary of the research design and no indication of which comments came from which treatment groups. Finally, though it is conceivable that some student participants may have had a pre-disposition towards certain strategies, this also is unlikely to have been a major risk. All participants did opt for traditional face-to-face instruction for their undergraduate

educations, after all, so it is possible that they might have shown some bias against online instruction, but it seems doubtful that they would have attempted to unduly influence the experiment as a result of some allegiance to a certain method of instruction.

Chapter Summary

This study has followed a between-factor experimental design in which 425 undergraduate students (of whom 303 actually participated in the study) were randomly assigned to one of five treatment groups in order to measure possible differences in key student outcomes across online instructional strategies. Differences between participating subjects were controlled for by the random assignment of subjects to treatment groups and by the inclusion of demographic factors in the regressions, so no pre-test was needed. Each group of participants was exposed to just one treatment, namely, the form of instructional strategy used to teach a single lesson. In addition to a control group which received instruction in a traditional classroom setting, the other four treatments represented the major paradigms in online instructional strategies. Aside from the strategy of instruction employed, other important variables such as the subject of the lesson, the instructor, discussion prompts, and the timing of the activities, were all held as constant as possible across the groups. Following the delivery of the lessons, each group was evaluated using the same instruments to determine their average scores on a test of comprehension, their average ratings on a survey of satisfaction, their frequency of participation and average quality of discussion contributions, and their rates of attrition throughout the experiment. Differences in these measures across groups were identified using ANOVA and chi-square tests, as well as linear and logistic regressions.

Chapter 4: Results

Organization of the Chapter

This chapter will present the results of the data analysis and findings from the study. It begins with a brief introduction outlining the purpose and rationale for how the data was analyzed. A synopsis of the execution of the study from recruitment through data collection follows. Next, the chapter summarizes and discusses the demographic information collected by the pre-lesson questionnaire. This is followed by an overview of the descriptive statistics pertaining to all of the outcome measures, coupled with a summary of the tests for normality for each numeric dependent variable. All of this leads to a presentation of the results from ANOVA and chi-square tests which compared the treatment group means for each of the dependent variables. These results are presented in separate sections, beginning with the results of an analysis of post-lesson assessment scores, followed by measures of class participation, then results from a survey of satisfaction, and finally, likelihood of study completion. The penultimate section of the chapter discusses the results of multiple linear regressions which incorporate both treatment groups and demographic characteristics as independent variables across each of the same ten outcome measures. Finally, the chapter concludes with a discussion of the tests for interaction effects between treatments and demographic variables.

Introduction

As previously noted, the purpose of this study was to investigate whether different strategies of online education may be associated with higher or lower measures of several important student outcomes for a single lesson. To this end, the treatment groups (each

associated with a different strategy of instruction), to which all participants were randomly assigned, was the key independent variable in the analyses of all dependent variables that were examined. Across all analyses, the group that received traditional, face-to-face instruction (Group C) served as the control and was therefore designated as the reference group variable for dummy coding. For each dependent variable, ANOVA or chi-square tests (depending on whether the dependent variable in question was numeric or dichotomous, respectively) were first used to examine whether one or more significant relationships existed between any of the independent variables and the dependent variable. Regressions (linear or logistic) were then employed in order to test the effects of demographic predictor variables on the outcome measures. Finally, new regressions were run which included categorical variables for the treatment groups and the demographic variables, in addition to a number of interaction terms represented by the product of each treatment group variable and one of the demographic variables. These last regressions were performed in order to investigate whether the employed strategy of instruction might have a more or less pronounced relationship to the outcome measures depending on whether the student participant belonged to a certain demographic category.

Execution of the Study

Nearly all aspects of the execution of the study went according to expectations. Recruitment of undergraduate student participants, which was considerably more difficult in the pilot study, was relatively easily accomplished within eight days of posting recruitment flyers on campus, to university-affiliated social networking sites, in the school newspaper, and to research Listservs. The relative ease of recruitment when compared to the pilot study was

most likely due to better distribution of recruiting materials and, probably most importantly, to doubling the participant compensation from $15 to $30. All registrations were taken on a first-come, first-serve basis.

As part of the registration process, all prospective participants were required to fill out a background demographic information survey (Appendix A). Since all fields within this study were coded as required for submission, there was no missing information from any of the 425 registrants. The registration process also included steps for participants to verify their status as undergraduate students at the university and their availability on the afternoon when the live events associated with Groups C, S, and H were to take place. When treatment groups were assignments and communicated to the participants (the same day), some participants who had been assigned to one of the three groups with live events asked to switch groups due to scheduling conflicts that had arisen in the intervening eleven days between the registration process and the communication of group assignments. All of these requests were denied, partly to preserve the integrity of the randomization of group assignments, and partly because potential scheduling conflicts are a natural factor in the dependent variable of study completion which was to be measured.

Treatment group assignments were designated and communicated to the participants one week before the live events associated with Groups C, S, and H were to take place. As soon as group assignments were emailed, all groups received access to their respective online group websites for the lesson. At the time when participants were granted access, the lesson websites for Groups C and S contained only instructions for participation in the study in addition to information about when and how to access their group's live events. In addition to

the instructions for their respective groups, the websites for the members of Groups H, V, and T also granted immediate access to the respective group's online instructional materials and, for groups V and T, access to instructor-proctored online discussion forums.

Some students in the groups with asynchronous online materials (Groups H, V, and T) opted to view their lesson materials and to post to the discussion forums immediately, while others waited until the very end of the week to engage with any of the content. Additionally, some of the participants assigned to these groups viewed the online materials only partially, while others accessed the materials multiple times. In addition, it can be presumed that the time and place, in addition to the amount of distractions present while viewing lesson materials, varied widely across participants in the asynchronous groups. Although differences in participants' interaction with online instructional materials undoubtedly affected their outcome measures, these differences, as previously noted, do not affect the objectives of the study because opportunities for students to engage with their learning materials differently are a realistic reflection of the differences between the instructional strategies being tested herein. If, however, a participant never accessed the online instructional materials, at the end of the week that student was prevented from accessing the post-lesson quiz and survey. Likewise, students in Groups C and S who did not attend their respective live lessons (in person or virtually) were also prevented from accessing the post-lesson quiz and survey, though they were counted as having attended even if they arrived late.

The live events for Groups C, S, and H all occurred on the same Friday, one week after the treatment groups were assigned. Participants who attended the live events for these groups signed in their attendance (including members of Group S in the videoconference) and

had their live sessions video-recorded. Unfortunately, participants in Groups C and H were not identified, except by voice and image, in the recordings, so it was not possible to match their comments to any demographic information obtained about them from the pre-survey. Although some students showed up or logged in late for their events, in no case did the late arrivals seriously disrupt the lessons or discussions. Also, although they were encouraged to log in early to work out any technical issues before the start of the lesson, some participants in Group S were still configuring their audio setups when the lesson began. Nevertheless, these minor technical issues were all resolved within the first five minutes of the lesson. It is also worth mentioning that the video image of the instructor (but not audio or slides) was twice interrupted during Group S's lesson when the instructor's screen-saver activated. These interruptions did not seriously affect the pace of the lesson, and may be interpreted as the type of unexpected, but not atypical, challenge inherent in synchronous virtual instruction. Also, in Groups C and S, students naturally asked questions and made comments during the lecture, which slowed down the pace of the lectures by about 5 minutes in both cases. All told, while there were many minor differences in delivery of the lessons and conduct of the discussions across groups, the majority of these discrepancies resulted from the application of different strategies of instruction, as opposed to notable differences in the execution of the lessons.

At the end of the Friday when the live events occurred, members of Groups H, V, and T lost access to their online instructional materials and discussion forums. At the same time, all participants gained access to the post-lesson quiz and survey of satisfaction. Both the quiz and survey could be accessed at any time during the week following the live events, though the quiz had a 30-minute time limit from when it was first accessed by each participant. Anticipating

that some students might view completion of the quiz as the only requirement for completion of the lesson, the treatment group websites were programmed so that the quiz could only be accessed after completion of the survey. Also, as with the pre-lesson demographic questionnaire, all questions for the post-lesson survey were coded as required, so it was impossible for participants to skip questions. If a participant left a question blank on the quiz they simply received a zero for that question.

Demographic Information

As noted above, completion of the pre-study background survey (Appendix A) was a requirement for registration in the study, with all questions mandatory, so complete data was obtained from all 425 students who registered for the study. Except when examining the dependent variable of study completion, however, only data from the 303 participants who completed the study were used in subsequent data analysis. Table 4.1 lists the demographic information collected from the pre-lesson survey separated by all registrants who were assigned to treatment groups versus those participants who actually completed the study.

Table 4.1. Demographic Information of Study Registrants and Participants

	Registrants		Participants	
	Total #s	Percentages	Total #s	Percentages
Total #	425		303	
Age				
<18	186	44%	135	45%
18-19	160	38%	113	37%
>19	79	19%	55	18%
Gender				
Male	113	27%	83	27%
Female	312	73%	220	73%
Race				
White	105	25%	68	22%
Asian	250	59%	184	61%
URM	70	16%	51	17%
1st Language				
English	282	66%	204	67%
Other	143	34%	99	33%
Major				
non-STEM	212	50%	149	49%
STEM	213	50%	154	51%
GPA				
<3.00	67	16%	49	16%
3.00-3.49	147	35%	103	34%
>3.50	211	50%	151	50%
Online Experience				
No	238	56%	179	59%
Yes	187	44%	124	41%
IP Law Experience				
No	386	91%	276	91%
Yes	39	9%	27	9%
Interest in Subject				
None/slight	145	34%	100	33%
Moderate+	280	66%	203	67%

Differences in demographic characteristics between the populations of registrants and participants will be examined in the section of this chapter that explores the outcome of study completion. For both registrants and participants, however, it is clear that there are notable differences between the students who took part in this study compared to the undergraduate population at the university where the study was conducted. In particular, female students made up 73% of the study registrants, as compared to 56% for the university population. Also, while Asian students do make up the largest racial group at the university, the proportion of study registrants who are of Asian descent is far greater than the percentage of Asian undergraduates at the university. Accordingly, the Caucasian and Under-Represented Minority students at the university were therefore underrepresented in the study. In addition, because only one-eighth of students at the university are international, it seems likely that the approximate one-third of study registrants who learned English as a second language also indicates an overrepresentation. Lastly, with over 80% of those who registered for the study listing their ages as nineteen or younger, it appears that most of the students who took part in the study were underclassmen.

It is difficult to know what caused certain groups at the university to be over- or under-represented in the study. Recruitment flyers, for the most part, were posted only in very general locations, such as the student union, and digital announcements were only posted to very generic forums, such as class year Facebook pages. More than likely, the interest in paid research studies is stronger with certain student groups and/or those groups are overrepresented in the research study Listservs that were used. Chapter 5 will include a discussion of the implications for the study results due to differences between the study and

university demographics. That section will also examine the university's representativeness of the entire population of U.S. undergraduate students and implications for the study related to those differences as well.

Overview of Outcome Measures

Because this study examined ten outcome measures related to student success, it seems advisable to review those measures and to summarize their descriptive statistics before delving into more in-depth analyses for each of the individual outcomes in turn. There were four categories of outcome measures in this study: participation/attrition, comprehension of instructional material, engagement in discussions, and satisfaction. Participation was measured by a single, dichotomous variable for whether a registrant completed the study. Comprehension was measured by each participant's total score on a quiz related to the instructional content of the lesson (Appendix B). Level of engagement in discussions was measured by two variables: a dichotomous variable for whether a participant made any comment in the lesson or discussion, and a numeric variable for each comment's quality score assigned by an outside grader according to a rubric of comment quality (Appendix D). Satisfaction was measured by a survey (Appendix C) administered to the participants after the lesson but before the quiz. Participants' satisfaction was broken down into six variables: one for their overall average score on the survey, and one average for each of five three-question sub-categories within the survey. These categories included questions to measure the participants' satisfaction with the subject matter, the instructor, the lesson format, their interactions with the instructor and with other students, and their overall learning experience.

The descriptive statistics for these variables, as well as the averages for each question on the survey of satisfaction, are presented in Table 4.2, broken down by treatment groups and with totals included for the entire study.

Table 4.2. Descriptive Statistics of Student Outcomes by Treatment

	Group C	Group S	Group H	Group V	Group T	Total
Overall Participation						
Assigned Registrants	85	85	85	85	85	425
Completing Participants	59	61	47	65	71	303
Percentage Completing	69.41%	71.76%	55.29%	76.47%	83.53%	71.29%
Assessment of Comprehension						
Assessment Score (out of 100)	77.06	59.67	70.64	63.39	69.58	67.88
Assessment Standard Deviation	17.17	18.6	15.14	17.29	16.89	18.05
Contributions to Class Discussion						
Total Comments	20	68	21	138	129	376
Participation Rate	30.51%	50.82%	40.43%	72.31%	54.93%	50.83%
Average Score per Comment	2.10	1.26	2.48	2.62	2.51	2.30
Comment Score Standard Dev.	0.72	0.56	0.60	0.84	0.85	0.93
Survey Averages (1-5 Scale)						
Overall Survey Average	3.70	3.48	3.70	3.26	3.13	3.43
Overall Survey Standard Deviation	0.44	0.54	0.50	0.50	0.57	0.56
Subject Matter Average	3.02	3.04	3.18	2.83	2.84	2.97
Subject Matter Std. Deviation	0.57	0.57	0.56	0.50	0.63	0.58
Q1: Subject Interest	3.53	3.44	3.65	3.08	3.01	3.31
Q2: Subject Challenge	2.39	2.45	2.57	2.45	2.50	2.47
Q3: Subject Usefulness	3.15	3.23	3.31	2.97	3.00	3.12
Instructor Average	4.25	3.97	4.22	3.81	3.49	3.92
Instructor Standard Deviation	0.55	0.46	0.55	0.44	0.57	0.58
Q4: Instructor Knowledgeable	4.63	4.47	4.73	4.42	3.93	4.41
Q5: Instructor Well-Organized	4.22	3.95	4.37	4.14	3.92	4.10
Q6: Instructor Engaging	3.90	3.50	3.55	2.88	2.63	3.24
Format Average	3.85	3.48	3.68	3.31	3.28	3.50
Format Standard Deviation	0.63	0.74	0.80	0.74	0.79	0.77
Q7: Format Appropriate	4.02	3.44	3.88	3.46	3.46	3.64
Q8: Format Captivating	3.68	3.39	3.37	2.86	2.68	3.17
Q9: Format Straightforward	3.85	3.60	3.78	3.6	3.69	3.70
Interactions Average	3.67	3.57	3.90	3.04	3.09	3.41
Interactions Standard Deviation	0.59	0.82	0.67	0.77	0.88	0.82
Q10: Instructor Interactions	3.81	3.42	3.86	2.72	2.74	3.26
Q11: Classmate Interactions	2.88	3.21	3.49	2.77	2.78	3.00
Q12: Questions Answered	4.31	4.08	4.35	3.62	3.75	3.99
Learning Experience Average	3.72	3.34	3.57	3.29	2.96	3.35
Learning Experience Standard Dev.	0.69	0.92	0.71	0.90	0.85	0.86
Q13: Learn More Than Normal	3.73	3.19	3.82	3.46	2.97	3.41
Q14: Interest in More of Same	3.46	3.18	3.10	3.00	2.67	3.10
Q15: Overall Satisfaction	3.98	3.65	3.80	3.40	3.24	3.59

In addition, while the sample size for each treatment was large enough (n>30) to assume a normal distribution of sample means (Agresti & Finlay, 2009, p. 94), it is also useful to examine the data in the sample itself for normality. The results of running Lilliefors normality test on the sample for each of the relevant dependent variables are listed below.

Table 4.3. Tests of Normality

Dependent Variable Abbreviation	Type of Variable	Lilliefors Test Statistic
Participation	Dichotomous	N/A
Post-Lesson Assessment of Comprehension	Numeric	0.120
Commenting (discussion participation)	Dichotomous	N/A
Average Score per Comment	Numeric	0.202
Satisfaction with Subject Matter	Numeric	0.117
Satisfaction with Instructor	Numeric	0.124
Satisfaction with Lesson Format	Numeric	0.117
Satisfaction with Interactions	Numeric	0.111
Satisfaction with Learning Experience	Numeric	0.102
Overall Survey Average	Numeric	0.068

For obvious reasons, the two dichotomous dependent variables were not tested for normality. For the other dependent variables, Lilliefors test did not indicate enough evidence at alpha=0.05 to reject the null hypothesis that the data are normally distributed. While this test does not confirm normality for any of sample outcome measures, it does allow for normality to be assumed in analyzing each of the outcome measures and in making comparisons between treatment groups using strategies such as analysis of variance.

Comprehension

The first outcome analyzed by this study, comprehension, was measured by each participant's score on a quiz that was administered following the lesson and discussion. All groups took the same quiz, which was developed from questions that the instructor regularly uses to assess comprehension in credit-bearing classes that he teaches at the university. The quiz, located in Appendix B, consisted of fifteen equally-weighted questions, and was scored out of a total of 100 points, or 6.67 points per question. There was a 30-minute time limit to complete the quiz and partial credit was not assigned on any questions. The descriptive statistics for the quiz are listed in Table 4.4.

Table 4.4. Quiz Score by Group

	Group C	Group S	Group H	Group V	Group T	Total
Completing Participants	59	61	47	65	71	303
Assessment Score (out of 100)	77.06	59.67	70.64	63.39	69.58	67.88
Assessment Standard Deviation	17.17	18.6	15.14	17.29	16.89	18.05

These results indicate an effect size of 0.963 and statistical power of >0.999. When one-way ANOVA was run with the quiz scores separated by treatment groups, the p-value of the F-statistic was <0.001, indicating that there is strong evidence for a significant relationship

between treatment group and mean quiz scores. Table 4.5 shows the results of this ANOVA with comparisons made across all treatment groups using the Bonferroni method to correct for the number of comparisons being made.

Table 4.5. ANOVA for Quiz Scores

C.I. Tests	Difference of Means	Bonferroni Lower	Bonferroni Upper	Bonferroni p-values
C-S: control-webinar	17.389	8.541	26.237	0.000
C-H: control-hybrid	6.423	-3.051	15.897	0.182
C-V: control-video	13.676	4.963	22.389	0.002
C-T: control-slides	7.485	-1.052	16.021	0.085
S-H: webinar-hybrid	-10.967	-20.371	-1.562	0.023
S-V: webinar-video	-3.713	-12.351	4.925	0.396
S-T: webinar-slides	-9.905	-18.364	-1.445	0.022
H-V: hybrid-video	7.253	-2.025	16.531	0.124
H-T: hybrid-slides	1.062	-8.050	10.174	0.818
V-T: video-slides	-6.191	-14.510	2.127	0.143

Note. All confidence intervals are at 95%

Comparing quiz scores between treatments shows that Group S (videoconference) and Group V (asynchronous video) averaged significantly lower scores on the quiz than the classroom students in Group C. In addition, the ANOVA suggests that Group S also scored significantly lower on the quiz overall than both Groups H (hybrid) and Group T (asynchronous slides). The fact that Group S scored lowest overall on the quiz is an interesting result because it suggests that students in the videoconference may have been more distracted than the other groups or may have otherwise found it more difficult to learn in that environment.

Class Participation

Likelihood of Participation

Class participation was measured by both the percentage of participants contributing via comment or question during the lesson or discussion, and by average scores of comment quality. The former measure was represented as a dichotomous variable scored as one if the participant made any comment and zero if he or she did not. For Groups C and H, both of which met in person, participation was counted if the participant made a relevant comment to the professor or to the rest of the class during the scheduled lesson or discussion. For Group S, comments were typed into a chat window during the videoconference. Comments made by participants in Groups V and T were collected from the discussion forums posted to their online group lesson pages.

Not every comment was counted as class participation. Interjections such as asking for the sign-in sheet or noting difficulty in hearing the instructor were not deemed as relevant to the lesson and therefore were not counted as class participation. That said, any comments or questions even remotely related to the content of the lesson or the discussion were counted as participation. Also, for obvious reasons, multiple comments made by a single participant did not affect the participation rate beyond the first comment. It should be pointed out that, across all five treatment groups, participation was encouraged, but was never described as mandatory. Nevertheless, it seems likely that, because discussion forums were posted to their lesson websites in Groups V and T, more students in those groups may have felt that commenting was expected or mandatory. Lastly, it is also worth noting that, in ways consistent with the different strategies that they were exposed to depending on their group assignments,

participants had very different time constraints upon their abilities to make comments. Participants in Groups C and S could make comments during single hour-long windows, fifteen minutes of which were specifically devoted to discussion. Members of Group H had just one twenty-minute window in which to participate. On the other hand, participants in Groups V and T could post their comments at any time during the span of an entire week. This was almost certainly the largest factor in the large difference in participation rates between the groups, as shown in Table 4.6.

Table 4.6. Participation Rates

	Group C	Group S	Group H	Group V	Group T	Total
Completing Participants	59	61	47	65	71	303
Participants who Commented	18	31	19	47	39	154
Participation Rate	30.51%	50.82%	40.43%	72.31%	54.93%	50.83%

These results indicate an effect size of 0.835 and statistical power of >0.999. Because participation is a dichotomous variable, it was appropriate to run chi-square tests to look for relationships between treatment groups. First, a chi-square test which included the participation outcomes from all five treatment groups yielded a chi-square statistic of 24.259 and a p-value of <0.001. Next, to determine which groups had significant differences in rates of participation, chi-square tests between paired groups were used. In these tests, the Bonferroni correction was used to account for the number of between-group comparisons being made. The results of the chi-square tests for differences in participation rates are shown in Table 4.7.

Table 4.7. Chi-Square Tests for Participation Rates

Comparison Groups	Difference in %	Chi-Square Stat	p-value	Significance
C-S: control-webinar	-20.311%	5.121	0.024	None
C-H: control-hybrid	-9.917%	1.132	0.287	None
C-V: control-video	-41.799%	21.665	0.000	C<V
C-T: control-slides	-24.421%	7.805	0.005	None
S-H: webinar-hybrid	10.394%	1.154	0.283	None
S-V: webinar-video	-21.488%	6.161	0.013	None
S-T: webinar-slides	-4.110%	0.223	0.637	None
H-V: hybrid-video	-31.882%	11.456	0.001	H<V
H-T: hybrid-slides	-14.504%	2.380	0.123	None
V-T: video-slides	17.378%	4.408	0.036	None

Note. Significance is determined at 95% confidence with the Bonferroni correction
Using the Bonferroni correction, the alpha for 95% confidence is 0.05/10 = 0.005

Table 4.7 indicates that the asynchronous online treatment Group V had a significantly higher rate of participation than the classroom and hybrid groups (C and H). These are not surprising results, given the additional time that the asynchronous groups had to write comments and the lack of constraints on simultaneous commenting in asynchronous discussion forums. Group S, which did not show evidence of being significantly different than any of the other groups in terms of participation rates, occupies something of a middle ground since students in a videoconference have the same overall time constraints placed on their comments but have no restrictions on simultaneous commenting.

Comment Quality

Every relevant comment (including multiple comments by the same participant) was also scored for quality. This process began with the transcription of all relevant comments from the recordings and forums of each treatment group. These transcribed comments were entered into a table with all identifying information, including treatment group, removed. Only

a brief note providing context to the circumstances under which the comment was made accompanied each comment. An outside grader, familiar with the subject matter, then used the table and a rubric for scoring discussion contributions (Appendix D) to attach a score from 1-4 to each comment. The averages from these scores are presented in Table 4.8.

Table 4.8. Comment Quality

	Group C	Group S	Group H	Group V	Group T	Total
Total Comments	20	68	21	138	129	376
Average Comment Score	2.10	1.26	2.48	2.62	2.51	2.30
Score Standard Deviation	0.72	0.56	0.60	0.84	0.85	0.93

The effect size and statistical power for comment scores based on these numbers are 1.460 and >0.999, respectively. ANOVA was then employed to investigate which instructional strategies had significantly higher or lower average comment scores. The results of this ANOVA are presented in Table 4.9.

Table 4.9. ANOVA for Comment Scores

	Difference	Bonferroni		Bonferroni
C.I. Tests	of Means	Lower	Upper	p-values
C-S: control-webinar	0.835	0.273	1.397	0.004
C-H: control-hybrid	-0.376	-1.067	0.314	0.277
C-V: control-video	-0.516	-1.045	0.013	0.056
C-T: control-slides	-0.412	-0.943	0.119	0.128
S-H: webinar-hybrid	-1.211	-1.763	-0.660	0.000
S-V: webinar-video	-1.351	-1.679	-1.024	0.000
S-T: webinar-slides	-1.247	-1.578	-0.916	0.000
H-V: hybrid-video	-0.140	-0.657	0.378	0.595
H-T: hybrid-slides	-0.035	-0.555	0.485	0.893
V-T: video-slides	0.104	-0.166	0.375	0.449

Note. All confidence intervals are at 95%

Unsurprisingly, these ANOVA results do not indicate a significant difference between the two groups that carried out discussions in person (C and H). On the other hand, the ANOVA

shows strong evidence that the videoconferenced treatment (Group S) was associated with lower overall scores on discussion comments compared to all other groups. It is also worth noting that the average length of comments varied considerably between the groups. In fact, the average word count per comment was 48 words for Group V, 44 words for Group T, 16 words for Group H, 10 words for Group C, and 6 words for Group S. The low word count for comments made by members of Group S may provide clues as to why their scores were rated lower on average. It may be that the chatroom format of the videoconference discussion discourages longer comments and, as a result, those comments may tend to lack some of the thoughtfulness and preparation of comments made in class or in discussion forums.

Satisfaction

The measure of the participants' satisfaction with their learning experiences was broken down into six dependent variables, all stemming from scores on an online survey (Appendix C) which the participants filled out after completing the lesson and discussion. Survey questions were broken down into sets of three questions each based on the following categories: participants' perception of the lesson's subject; participants' perception of the instructor; format of the lesson; participants' feelings about opportunities for interaction during the lesson; and participants' satisfaction with the overall learning experience. An additional dependent variable was included to measure each participant's average score across the entire survey. All fifteen questions in the survey were scored from one through five, with scores of five being the most desirable. Because all questions in the survey were coded as mandatory, no survey could be submitted with missing information.

Full Survey Average

Each study participant was assigned one measure for satisfaction that was determined by averaging their scores across all fifteen survey questions. The averages and standard deviations for the full survey by treatment group are listed in Table 4.10.

Table 4.10. Full Survey Averages

	Group C	Group S	Group H	Group V	Group T	Total
Total Surveys	59	61	47	65	71	303
Average Score per Item	3.70	3.48	3.70	3.26	3.13	3.43
Standard Deviation	0.44	0.54	0.50	0.50	0.57	0.56

The effect size for the average scores on the full survey was 1.020, yielding a statistical power calculation of >0.999. Running a single-factor ANOVA across the treatment groups provided a p-value of <0.001 for the F-statistic. The results of this ANOVA are listed in Table 4.11.

Table 4.11. ANOVA for Survey of Satisfaction

	Difference	Bonferroni		Bonferroni
C.I. Tests	of Means	Lower	Upper	p-values
C-S: control-webinar	0.228	-0.037	0.494	0.091
C-H: control-hybrid	0.004	-0.281	0.288	0.979
C-V: control-video	0.447	0.186	0.709	0.001
C-T: control-slides	0.557	0.301	0.813	0.000
S-H: webinar-hybrid	-0.225	-0.507	0.058	0.118
S-V: webinar-video	0.219	-0.041	0.478	0.097
S-T: webinar-slides	0.329	0.075	0.583	0.012
H-V: hybrid-video	0.444	0.165	0.722	0.002
H-T: hybrid-slides	0.553	0.280	0.827	0.000
V-T: video-slides	0.110	-0.140	0.360	0.386

Note. All confidence intervals are at 95%

As the ANOVA indicates, both of the online asynchronous groups (V and T) tended to provide significantly lower scores on the survey of satisfaction compared to the other groups,

all of which contained synchronous elements and which were not significantly different from each other. This seems to indicate that the participants in this study may value direct, real-time interaction and that they were less satisfied with methods of instruction in which that live element was lacking.

Subject Matter Satisfaction

The first three questions from the post-lesson survey concerned the participants' satisfaction with the subject matter for the lesson. While the subject of the lesson for each treatment group was obviously the same, the question of whether students would have a different perception of the subject depending on which instructional strategy they received was deemed to be worth exploring. The three questions pertaining to subject asked participants to gauge how interesting, challenging, and useful they found the subject to be. The average scores on each of these three questions, as well as the averages of all three, are displayed in Table 4.12.

Table 4.12. Survey Subject Averages

	Group C	Group S	Group H	Group V	Group T	Total
Total Surveys	59	61	47	65	71	303
Q1: Subject Interest	3.53	3.44	3.65	3.08	3.01	3.31
Q2: Subject Challenge	2.39	2.45	2.57	2.45	2.50	2.47
Q3: Subject Usefulness	3.15	3.23	3.31	2.97	3.00	3.12
Subject Score Average	3.02	3.04	3.18	2.83	2.84	2.97
Subject Score Std. Deviation	0.57	0.57	0.56	0.50	0.63	0.58

From the above table, it is worth noting that there does not seem not be much difference across treatment groups in terms of how challenging the participants found the subject. There were, however, considerable differences between how interesting and how useful participants perceived the subject to be. Regardless, effect size and statistical power

were analyzed for the average of the scores on these three subject-related survey questions, providing results of 0.592 and >0.999, respectively. Comparing the results for the three-question average across treatment groups using ANOVA provided a p-value of 0.007. The results from this ANOVA are listed in Table 4.13.

Table 4.13. ANOVA for Subject Questions

C.I. Tests	Difference of Means	Bonferroni Lower	Bonferroni Upper	Bonferroni p-values
C-S: control-webinar	-0.005	-0.298	0.289	0.975
C-H: control-hybrid	-0.155	-0.469	0.159	0.331
C-V: control-video	0.192	-0.097	0.481	0.191
C-T: control-slides	0.173	-0.110	0.456	0.229
S-H: webinar-hybrid	-0.150	-0.462	0.162	0.343
S-V: webinar-video	0.197	-0.090	0.483	0.177
S-T: webinar-slides	0.178	-0.103	0.458	0.213
H-V: hybrid-video	0.347	0.039	0.654	0.028
H-T: hybrid-slides	0.328	0.025	0.630	0.034
V-T: video-slides	-0.019	-0.295	0.257	0.892

Note. All confidence intervals are at 95%

The ANOVA results from Table 4.13 indicate that both of the asynchronous groups (V and T) rated their overall perceptions of the lesson's subject matter as lower, on average, than did the hybrid group (Group H).

Satisfaction with the Instructor

Questions four through six on the post-lesson survey dealt with the students' perceptions of their instructor. As with the lesson's subject, the instructor was the same across all treatment groups, so differences between average scores given by participants in each group may provide evidence that the strategy of instruction influences how students perceive their instructors. It should be noted that only Groups C and H ever met with the instructor in person.

Students in Group S could view the instructor in a live video and, while students in Group V also saw the instructor via video, that video was pre-recorded. Members of Group T never even saw a picture of the instructor—only his slides and written comments. The three questions comprising this section of the survey asked participants to rate how knowledgeable, well-organized, and engaging the instructor seemed to them. The average scores on each of these questions, in addition to the averages across all three questions, are presented in Table 4.14.

Table 4.14. Survey Instructor Averages

	Group C	Group S	Group H	Group V	Group T	Total
Total Surveys	59	61	47	65	71	303
Q4: Instructor Knowledgeable	4.63	4.47	4.73	4.42	3.93	4.41
Q5: Instructor Well-Organized	4.22	3.95	4.37	4.14	3.92	4.10
Q6: Instructor Engaging	3.90	3.50	3.55	2.88	2.63	3.24
Instructor Average	4.25	3.97	4.22	3.81	3.49	3.92
Instructor Standard Deviation	0.55	0.46	0.55	0.44	0.57	0.58

The effect size for the questions pertaining to perceptions of the instructor was 1.305, yielding a statistical power of >0.999. Next, the results of an ANOVA comparing scores for instructor satisfaction across the treatment groups are presented in Table 4.15.

Table 4.15. ANOVA for Instructor Questions

	Difference	Bonferroni		Bonferroni
C.I. Tests	of Means	Lower	Upper	p-values
C-S: control-webinar	0.270	0.004	0.537	0.047
C-H: control-hybrid	0.057	-0.229	0.343	0.693
C-V: control-video	0.438	0.176	0.701	0.001
C-T: control-slides	0.746	0.489	1.004	0.000
S-H: webinar-hybrid	-0.213	-0.497	0.070	0.139
S-V: webinar-video	0.168	-0.093	0.428	0.204
S-T: webinar-slides	0.476	0.221	0.731	0.000
H-V: hybrid-video	0.381	0.101	0.661	0.008
H-T: hybrid-slides	0.689	0.414	0.964	0.000
V-T: video-slides	0.308	0.057	0.559	0.017

Note. All confidence intervals are at 95%

These results indicate that all of the fully-online groups (S, V, and T) formed a less-favorable view of the instructor compared to the students in Group C, who had their lesson in person. Group H also gave the instructor significantly higher ratings than either of the asynchronous groups (V and T). Unsurprisingly, Group T, whose members never saw the instructor at all, gave him the lowest scores, even compared to Group V, who only saw him in a pre-recorded video.

Format of the Lesson

Students' perceptions of the format of the lesson, or the instructional strategy employed, is obviously highly relevant to this study. Questions seven through nine on the survey asked the participants to what extent they believed that the format of the lesson was appropriate to the subject of the lesson; to what extent the format of the lesson created a captivating learning experience; and to what extent the format of the lesson made it easy for them to follow along and fulfill their obligations as students. It is worth noting here that this was one area of the survey in which the subject pool from which the participants for this study were drawn may have had a considerable influence on the results, since all of the subjects were students who elected to complete their undergraduate studies in a traditional face-to-face learning environment. Therefore, it is not at all surprising to see the highest scores for the control Group C across all three questions, as shown in Table 4.16.

Table 4.16. Survey Format Averages

	Group C	Group S	Group H	Group V	Group T	Total
Total Surveys	59	61	47	65	71	303
Q7: Format Appropriate	4.02	3.44	3.88	3.46	3.46	3.64
Q8: Format Captivating	3.68	3.39	3.37	2.86	2.68	3.17
Q9: Format Straightforward	3.85	3.60	3.78	3.6	3.69	3.70
Format Average	3.85	3.48	3.68	3.31	3.28	3.50
Format Standard Deviation	0.63	0.74	0.80	0.74	0.79	0.77

One important observation from the numbers in Table 4.16 is that there was comparatively little difference in terms of how members in each of the five treatment groups felt regarding their respective format's ease of use. The differences in the three-question average scores for format were largely due to the other two questions, particularly Question 8, which asked to what extent the format of the lesson created a captivating learning experience. Using the average for all three questions pertaining to format, the effect size was calculated at 0.745—which indicates a statistical power of >0.999. The results from a single-factor ANOVA using the average format scores separated by treatment groups are shown in Table 4.17.

Table 4.17. ANOVA for Format Questions

	Difference	Bonferroni		Bonferroni
C.I. Tests	of Means	Lower	Upper	p-values
C-S: control-webinar	0.361	-0.022	0.744	0.064
C-H: control-hybrid	0.188	-0.222	0.598	0.365
C-V: control-video	0.530	0.153	0.906	0.006
C-T: control-slides	0.556	0.187	0.926	0.003
S-H: webinar-hybrid	-0.173	-0.580	0.234	0.400
S-V: webinar-video	0.168	-0.205	0.542	0.374
S-T: webinar-slides	0.195	-0.171	0.561	0.293
H-V: hybrid-video	0.342	-0.060	0.743	0.094
H-T: hybrid-slides	0.368	-0.026	0.763	0.067
V-T: video-slides	0.027	-0.333	0.387	0.883

Note. All confidence intervals are at 95%

This ANOVA indicates that, as with several of the other measures of satisfaction, the online asynchronous groups (V and T) scored the survey questions about format significantly lower than the face-to-face group.

Opportunities for Interaction

Survey questions ten through twelve dealt with the participants' satisfaction with opportunities for interaction with the instructor, with classmates, and the extent to which their questions about the lesson were answered. The average scores on these questions are listed in Table 4.18.

Table 4.18. Survey Instructor Averages

	Group C	Group S	Group H	Group V	Group T	Total
Total Surveys	59	61	47	65	71	303
Q4: Instructor Knowledgeable	4.63	4.47	4.73	4.42	3.93	4.41
Q5: Instructor Well-Organized	4.22	3.95	4.37	4.14	3.92	4.10
Q6: Instructor Engaging	3.90	3.50	3.55	2.88	2.63	3.24
Instructor Average	4.25	3.97	4.22	3.81	3.49	3.92
Instructor Standard Deviation	0.55	0.46	0.55	0.44	0.57	0.58

The effect size and statistical power for the average scores on the three questions pertaining to interactions across treatment groups were 1.053 and >0.999, respectively. Table 4.19 presents the results of a one-way ANOVA comparing average scores for these three questions across the five treatment groups.

Table 4.19. ANOVA for Interaction Questions

C.I. Tests	Difference of Means	Bonferroni Lower	Bonferroni Upper	Bonferroni p-values
C-S: control-webinar	0.109	-0.284	0.503	0.584
C-H: control-hybrid	-0.234	-0.656	0.188	0.273
C-V: control-video	0.631	0.243	1.018	0.002
C-T: control-slides	0.559	0.179	0.939	0.004
S-H: webinar-hybrid	-0.343	-0.762	0.075	0.107
S-V: webinar-video	0.521	0.137	0.906	0.008
S-T: webinar-slides	0.449	0.073	0.826	0.020
H-V: hybrid-video	0.865	0.452	1.278	0.000
H-T: hybrid-slides	0.793	0.387	1.198	0.000
V-T: video-slides	-0.072	-0.442	0.298	0.701

Note. All confidence intervals are at 95%

Following a similar trend from other sub-sections of the survey, Table 4.19 indicates that participants in both asynchronous online groups (V and T) tended to rate the quality of their interactions as lower than those participants who received instruction in any other format. There was not enough evidence to determine a significant difference between any of the other treatment groups. Since both Groups V and T used discussion forums to carry out all of their interactions, these results seem to indicate that students may value live interactions (whether online or in person) more highly than asynchronous, text-based communication.

Learning Experience

The final three questions on the survey of satisfaction were more general in nature, and together served as a gauge for the participants' overall satisfaction with their learning experiences. Question 13 asked the participants to compare how much they learned in the study to a typical hour of class time spent in their undergraduate courses of study. Question 14 asked the participants how likely they would be to sign up for a full class on the subject if it

were to be offered by the same instructor using a similar format of instruction. Question 15 asked the participants to directly rate their satisfaction with the learning experience provided to them. The average scores on these questions across the treatment groups are presented in Table 4.20.

Table 4.20. Survey Learning Experience Averages

	Group C	Group S	Group H	Group V	Group T	Total
Total Surveys	59	61	47	65	71	303
Q13: Learn More Than Normal	3.73	3.19	3.82	3.46	2.97	3.41
Q14: Interest in More of Same	3.46	3.18	3.10	3.00	2.67	3.10
Q15: Overall Satisfaction	3.98	3.65	3.80	3.40	3.24	3.59
Learning Experience Average	3.72	3.34	3.57	3.29	2.96	3.35
Learning Experience Standard Dev.	0.69	0.92	0.71	0.90	0.85	0.86

For the average of these three questions, the effect size was determined to be 0.888, which provides a statistical power measurement of >0.999. Running ANOVA on these average scores separated by treatment groups provided the results shown in Table 4.21.

Table 4.21. ANOVA for Learning Experience Questions

	Difference	Bonferroni		Bonferroni
C.I. Tests	of Means	Lower	Upper	p-values
C-S: control-webinar	0.384	-0.042	0.811	0.077
C-H: control-hybrid	0.163	-0.294	0.620	0.481
C-V: control-video	0.436	0.016	0.856	0.042
C-T: control-slides	0.751	0.340	1.163	0.000
S-H: webinar-hybrid	-0.221	-0.675	0.232	0.335
S-V: webinar-video	0.052	-0.365	0.468	0.807
S-T: webinar-slides	0.367	-0.041	0.775	0.077
H-V: hybrid-video	0.273	-0.174	0.720	0.229
H-T: hybrid-slides	0.588	0.149	1.028	0.009
V-T: video-slides	0.315	-0.086	0.716	0.122

Note. All confidence intervals are at 95%

Consistent with other measures from the post-lesson survey of satisfaction, the results of the ANOVA above show that participants in Groups V and T tended to rate their overall learning experiences as poorer than the participants in Group C. This was also true for the participants in Group T when compared to Group H.

Study Completion/Attrition

The final dependent variable to be examined is the rate at which students who registered for the study actually completed it. This variable was coded as dichotomous for completion or non-completion. Participants were judged to have completed the study if, and only if, they completed the post-lesson quiz. There were, however, several hard-coded prerequisites for completing the quiz. The first of these was that, as part of the registration process, all students taking part in the study were required to complete the demographic questionnaire. Additionally, in order to gain access the post-lesson survey of satisfaction, participants needed to have at least accessed the learning materials for the study. For Groups C and S, this entailed attending or accessing the live components of their lessons. For Groups H, V, and T the standard of accessing the instructional content was less rigid; the group lesson website only needed to show evidence that the participant had clicked into the learning materials at least once. Finally, the post-lesson quiz could not be accessed without first completing the survey. The rates of study completion by treatment group are displayed in Table 4.22.

Table 4.22. Study Completion

	Group C	Group S	Group H	Group V	Group T	Total
Assigned Registrants	85	85	85	85	85	425
Completing Participants	59	61	47	65	71	303
Percentage Completing	69.41%	71.76%	55.29%	76.47%	83.53%	71.29%

Across treatment groups, the effect size for study completion was 0.623 and the resulting statistical power was calculated at >0.999. A preliminary chi-square test conducted using a contingency table which included study completion numbers for all five treatment groups yielded a chi-square statistic of 18.119 and a p-value of 0.001. Table 4.23 shows the results of chi-square tests comparing the rates of study completion between paired treatment groups.

Table 4.23. Chi-Square Tests for Study Completion Rates

Comparison Groups	Difference in %	Chi-Square Stat	p-value	Significance
C-S: control-webinar	-2.353%	0.113	0.736	None
C-H: control-hybrid	14.118%	3.609	0.057	None
C-V: control-video	-7.059%	1.073	0.300	None
C-T: control-slides	-14.118%	4.708	0.030	None
S-H: webinar-hybrid	16.471%	4.976	0.026	None
S-V: webinar-video	-4.706%	0.491	0.484	None
S-T: webinar-slides	-11.765%	3.389	0.066	None
H-V: hybrid-video	-21.176%	8.479	0.004	C<D
H-T: hybrid-slides	-28.235%	15.958	0.000	C<E
V-T: video-slides	-7.059%	1.324	0.250	None

Note. Significance is determined at 95% confidence with the Bonferroni correction
Using the Bonferroni correction, the alpha for 95% confidence is 0.05/10 = 0.005

This main takeaway from these comparisons is that members of Group H (hybrid) were less likely to complete the study when compared to either of the asynchronous treatment groups (V and T). Group H's lower completion rate is somewhat challenging to explain, since they had the exact same instructional content to watch as Group V. On the other hand, members of Group H may have seen their dual responsibilities of watching the video and attending a discussion in person as more onerous. Some of the participants assigned to Group H may also have felt worried about showing up to the discussion without having carefully

watched the video. Conversely, the higher completion rates for Groups V and T may be due to the greater scheduling flexibility which was afforded to them by not having a set time when they had to log in or attend class in person.

Regression Results

In addition ANOVAs and chi-square tests, this study also used regressions in order to both determine the predicted effects of the different instructional strategies compared to classroom instruction and to measure any effects of the demographic variables upon the various student outcomes. It should be reiterated that the primary objective of this study was to look for significant differences in the effects of different online instructional strategies on student outcomes—and this was satisfactorily accomplished using ANOVAs and chi-square tests presented thus far. On the other hand, while the multiple regression techniques used by this study only compared each online instructional strategy to the classroom control group, these regressions provided greater specificity in terms of measuring the online strategies' predicted effects upon the student outcome measures benchmarked against the classroom. Using regressions also allowed demographic variables to be added to the analysis, which is essential to understanding any possible interaction effects.

For each of the eight numeric dependent variables, multiple linear regressions were used to measure the predicted effects of each online instructional strategy compared to classroom instruction. For the two dichotomous outcomes—class participation and study completion—logistic regressions were used to examine the relationships between treatments and the likelihood of a student participating or finishing the study. Also, for the measures of class participation and comment quality it was not possible to match the participants who were

commenting to their demographic characteristics, so these regressions were run with only the treatment groups as independent variables. For all of the other outcome measures, the categorical variables for treatment groups (4 total) and demographic characteristics (12 total) were all included as independent variables in the regressions. The classroom control group (C) was used as the reference group in all regressions. The reference groups for each category of demographic variable are listed in Table 3.1. The ten multiple regressions (one for each outcome measure) are listed in Appendices E through N. A summary of the significant coefficients from these regressions are also shown in Table 4.24.

Table 4.24. Significant Coefficients from all Regressions at 95% Confidence

Predictor Variable	Assessment Score	Participation Rate	Comment Quality	Satisfaction: Overall	Satisfaction: Subject	Satisfaction: Instructor	Satisfaction: Format	Satisfaction: Interactions	Satisfaction: Experience	Study Completion
Group S	−	+	−	n.s.	n.s.	−	−	n.s.	n.s.	n.s.
Group H	−	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	−
Group V	−	+	+	−	−	−	−	−	−	n.s.
Group T	−	+	+	−	n.s.	−	−	−	−	+
18-19 years old	n.s.			n.s.	n.s.	n.s.	+	n.s.	n.s.	n.s.
>19 years old	n.s.			n.s.	n.s.	n.s.	n.s.	n.s.	−	n.s.
Female	−			n.s.	n.s.	n.s.	n.s.	n.s.	−	n.s.
Asian	n.s.			n.s.	+	n.s.	n.s.	n.s.	n.s.	n.s.
URM	n.s.			n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
English 2nd Language	−			+	+	n.s.	n.s.	n.s.	+	n.s.
STEM Major	n.s.			n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
3.00-3.49 GPA	n.s.			n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
>3.50 GPA	n.s.			n.s.	−	n.s.	n.s.	n.s.	n.s.	n.s.
Online Experience	n.s.			n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	−
IP Law Experience	n.s.			n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Moderate+ Interest	n.s.			n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

It should be pointed out that the regressions show some significant differences (at alpha=0.05) between the effects of the treatment group variables that were not represented in the ANOVA and chi-square tests. These differences are likely the result of the Bonferroni correction being applied to the ANOVA and chi-square tests, but not the tests for significance of the regression coefficients. Among the notable differences between the regressions versus the ANOVA and chi-square tests is the observation that the classroom control group (C) scored

significantly higher than all other treatment groups on the assessment of comprehension. In the ANOVA for assessment score results, the differences between the control group (C) and Groups H and T (hybrid and asynchronous text) were not significant. Another difference: in the regression for comment quality—but not the ANOVA—the two asynchronous groups (V and T) scored significantly higher than the control group. Finally, in the logistic regression for study completion, the hybrid group (H) was shown to be significantly less likely to complete the study compared to the control group (C), while the asynchronous group that used text and slides (T) was shown to be significantly more likely to complete—differences that were not significant in the chi-square test for that outcome. Again, the fact that these results were significant in the regressions, but not in the ANOVAs and chi-square test is probably due to the conservativism enforced by the Bonferroni correction being applied to the latter tests.

Among the interesting results pertaining to the demographic variables was the observation that gender and having learned English as a second language both showed significant differences at alpha=0.05 in the assessment of comprehension. In the case of English as a second language, it is not surprising that average quiz scores would be significantly lower for non-native English speakers since the language barrier may have accounted for some difficulty in understanding the lesson or the questions on the quiz. The indication that female participants in this study performed significantly lower on the quiz than male participants was much more unexpected and is of some concern. This scoring differential of almost a full question is unexplained and may raise some concerns about the design of the assessment, the delivery of the lesson across all treatment groups, or the study's methods of recruitment. With regards to the survey of satisfaction, participants who learned English as a second language

gave significantly higher ratings on the survey, on average, than did native English speakers. There are many possible explanations for this difference, such as the appeal of the instructor, the subject of the lesson, or factors unrelated to this specific lesson such as culture. Finally, for the outcome of study completion, it is surprising to note that the experience of having previously taken an online course correlated to a reduced rate of completion for the study. Overall, from the perspective of this study, the significant regression coefficients for these demographic variables are most interesting to note in terms of their potential for interactions with the different strategies of instruction.

Interaction Effects

To test for the presence of interaction effects between treatment and demographic variables, seventy-two new regressions were run, each with the purpose of testing for the possibility of an interaction between the treatments and a single demographic variable across eight of the outcome variables–excluding the two outcome variables related to class participation for which demographic information was not available. With four comparison groups for the treatment group variable and a total of twelve comparison groups across the nine demographic variables, this means that forty-eight interaction terms were tested for each of the eight dependent variables for which demographic information was available. The result is that a total of 384 possible interaction effects were tested for signficance.

Despite this large number of tests, only three of the regressions returned coefficients for any of the interaction terms that were significant at alpha=0.05. Some research indicates that, while seldom used, it may be appropriate to apply the Bonferroni correction when testing for the significance of independent variables in multiple linear regression (Mundfrom, Perrett,

Schaffer, Piccone, & Roozeboom, 2006). When the Bonferroni correction is applied to this study's tests for significant interaction effects, the alpha for each of the tests of significance becomes 0.05/384 = 0.00013, a level at which none of the coefficients for the interaction terms were significant. Keeping this in mind, it may still be informative to examine the three regressions which resulted in interaction terms that were significant at the uncorrected alpha=0.05. The first of these, which tested for the possibility of interaction effects between treatments and age in participants' post-lesson quiz scores, is presented below in Table 4.25.

Table 4.25. Quiz Score Regression with Interactions between Age & Treatment Group

Regression Statistics				
Observations	303		303	
Multiple R	0.430		0.479	
R Square	0.185		0.229	
Adjusted R Square	0.139		0.163	
	Without Interactions		With Interactions	
	Coefficients	p-values	Coefficients	p-values
Intercept	81.473	0.000	84.622	0.000
Group S	-15.887	0.000	-16.200	0.001
Group H	-7.036	0.037	-6.350	0.202
Group V	-12.711	0.000	-13.837	0.003
Group T	-6.897	0.022	-14.880	0.001
18-19 years old	2.675	0.230	-0.788	0.870
>19 years old	-0.308	0.915	-5.241	0.431
Female	-5.357	0.016	-5.635	0.012
Asian	-1.024	0.679	-1.060	0.667
URM	-1.682	0.606	-1.577	0.630
English 2nd Language	-5.903	0.007	-6.252	0.005
STEM Major	0.460	0.820	-0.294	0.884
3.00-3.49 GPA	-2.753	0.356	-4.009	0.182
>3.50 GPA	-0.937	0.745	-1.553	0.590
Online Experience	-2.324	0.259	-1.799	0.380
IP Law Experience	4.080	0.252	4.106	0.249
Moderate+ Interest	3.844	0.068	4.351	0.039
B*18-19			1.767	0.804
C*18-19			-1.561	0.827
D*18-19			3.143	0.632
E*18-19			10.391	0.114
B*>19			0.327	0.969
C*>19			-3.850	0.715
D*>19			-2.862	0.755
E*>19			23.105	0.008

As shown in the above table, the interaction terms between Group T and students who are twenty years old or older was significantly positive. Since, in the regression without interaction terms, Group T's treatment correlated with lower scores on the quiz while the effect

of age was not significant, the simplest interpretation of this interaction term is that the negative effect on quiz scores for students who received their lesson through the asynchronous strategy employed by Group T was not as pronounced for older students. In other words, the oldest students who were assigned to Group T tended to perform less poorly (relative to the face-to-face group) on the post-lesson quiz when compared to other Group T participants. Because the free navigation of annotated slides may have made it easiest for students in Group T to cut corners in viewing their lesson, it is possible that the older students who were assigned to this group were less likely to skim through. This could be because the older students were not as busy as first-year students or were simply better able to manage their schedules. Alternatively, the older students may have had more experience with self-directed learning from their advanced classes and may have found the format of instruction in Group T more familiar and easier to use.

The second interaction term that displayed results which were significant at alpha=0.05 was the term for interactions between non-native English speakers and Group H for the overall average on the survey of satisfaction. The results of this regression are displayed in Table 4.26.

Table 4.26. Full Survey Regression with Interactions between Language & Treatment

Regression Statistics				
Observations	303		303	
Multiple R	0.482		0.519	
R Square	0.232		0.269	
Adjusted R Square	0.189		0.217	
	Without Interactions		With Interactions	
	Coefficients	p-values	Coefficients	p-values
Intercept	3.575	0.000	3.579	0.000
Group S	-0.166	0.081	-0.091	0.424
Group H	0.011	0.914	-0.133	0.259
Group V	-0.440	0.000	-0.497	0.000
Group T	-0.546	0.000	-0.560	0.000
18-19 years old	0.132	0.050	0.151	0.023
>19 years old	-0.169	0.053	-0.165	0.056
Female	-0.012	0.857	0.005	0.945
Asian	0.028	0.703	0.029	0.697
URM	0.119	0.227	0.115	0.234
English 2nd Language	0.185	0.005	0.111	0.414
STEM Major	0.059	0.331	0.058	0.331
3.00-3.49 GPA	-0.039	0.663	-0.024	0.786
>3.50 GPA	-0.106	0.219	-0.112	0.188
Online Experience	0.038	0.543	0.024	0.696
IP Law Experience	-0.065	0.543	-0.042	0.688
Moderate+ Interest	0.040	0.523	0.046	0.465
B*English 2nd Language			-0.233	0.227
C*English 2nd Language			0.572	0.010
D*English 2nd Language			0.171	0.366
E*English 2nd Language			0.039	0.834

In Table 4.26, the presence of a positive coefficient for the interaction of English as a second language and Group H suggests that the effects of one of those variables was enhanced by the presence of the other. Examining the original regression reveals that Group H did not return results that were significantly different from Group C on the survey of satisfaction, while students who learned English as a second language actually rated their overall experiences in

the study significantly more positively than the other students who grew up speaking English. This suggests that, when non-native English speakers were in Group H, they tended to be even more positive about their experiences with the lesson. It is certainly possible that non-native English speakers appreciated the hybrid format of Group H, valuing both the opportunity to interact with the professor in person and the ability to pause or slow down the pre-recorded video lecture to review parts that might have involved difficult vocabulary and/or expressions.

The last regression to provide significant interaction coefficients included terms for possible interactions between the treatments and students' grade point averages in terms of their effects on how the participants rated the subject of the lesson in the post-lesson survey of satisfaction. The results of this regression are shown in Table 4.27.

Table 4.27. Subject Questions Regression with Interactions between GPA & Treatment

Regression Statistics				
Observations	303		303	
Multiple R	0.438		0.472	
R Square	0.192		0.223	
Adjusted R Square	0.147		0.156	
	Without Interactions		With Interactions	
	Coefficients	p-values	Coefficients	p-values
Intercept	2.674	0.000	2.349	0.000
Group S	0.033	0.744	0.180	0.468
Group H	0.172	0.108	0.632	0.010
Group V	-0.203	0.038	0.135	0.608
Group T	-0.187	0.050	0.395	0.083
18-19 years old	0.043	0.540	0.034	0.636
>19 years old	-0.145	0.116	-0.143	0.122
Female	0.102	0.148	0.063	0.383
Asian	0.282	0.000	0.302	0.000
URM	0.200	0.055	0.200	0.059
English 2nd Language	0.275	0.000	0.272	0.000
STEM Major	0.003	0.964	0.000	0.994
3.00-3.49 GPA	-0.112	0.240	0.340	0.117
>3.50 GPA	-0.193	0.036	0.169	0.390
Online Experience	0.128	0.052	0.137	0.038
IP Law Experience	-0.068	0.547	-0.089	0.433
Moderate+ Interest	0.109	0.104	0.140	0.040
B*3-3.49 GPA			-0.200	0.510
C*3-3.49 GPA			-0.587	0.056
D*3-3.49 GPA			-0.511	0.107
E*3-3.49 GPA			-0.809	0.005
B*>3.50 GPA			-0.177	0.532
C*>3.50 GPA			-0.571	0.050
D*>3.50 GPA			-0.352	0.234
E*>3.50 GPA			-0.648	0.015

This regression shows two significant interaction terms between GPA and the treatment for Group T, both of them negative. Since Group T students tended to rate the subject matter of the lesson as significantly lower compared to Group C in the original regression, these

negative coefficients for the interaction terms between Group T and the higher-GPA students suggests that higher-achieving students in Group T may have felt even more negatively about the subject matter compared to the other students assigned to Group T. This also implies that the lower-GPA students in Group T felt relatively less negatively about the subject. One possible explanation for this significant interaction term is that higher-GPA students tend to be more invested in their learning experience and may therefore have been more put off by what they may have perceived as a less-valuable or less-engaging teaching strategy in Group T's annotated slides.

Chapter Summary

This chapter has presented the results of an investigation into ten different student outcomes, each of which was measured across five different instructional strategies for a single lesson. Somewhat unexpectedly, when tested with ANOVA or chi-square, all ten of these outcomes showed evidence (significant at alpha=0.05) of differences between the groups that had been exposed to different strategies of instruction. A summary of some of the most important results from the ANOVAs and chi-square tests that were presented in this chapter are shown in Table 4.28.

Table 4.28. All ANOVA and Chi-Squared Results at 95% Confidence

Outcome Measure	C-S	C-H	C-V	C-T	S-H	S-V	S-T	H-V	H-T	V-T
Assessment Score	+	n.s.	+	n.s.	–	n.s.	–	n.s.	n.s.	n.s.
Participation Rate	n.s.	n.s.	–	n.s.	n.s.	n.s.	n.s.	–	n.s.	n.s.
Comment Quality	+	n.s.	n.s.	n.s.	–	–	–	n.s.	n.s.	n.s.
Satisfaction: Overall	n.s.	n.s.	+	+	n.s.	n.s.	+	+	+	n.s.
Satisfaction: Subject	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	+	+	n.s.
Satisfaction: Instructor	+	n.s.	+	+	n.s.	n.s.	+	+	+	+
Satisfaction: Format	n.s.	n.s.	+	+	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Satisfaction: Interaction	n.s.	n.s.	+	+	n.s.	+	+	+	+	n.s.
Satisfaction: Experience	n.s.	n.s.	+	+	n.s.	n.s.	n.s.	n.s.	+	n.s.
Study Completion	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	–	–	n.s.

The results from Table 4.28 were taken from between-group comparisons using the Bonferroni method of correcting for the number of comparisons and confidence levels of 95%. In Table 4.28 a "+" or "-" is used to indicate a significantly positive or negative relationship between groups, while the abbreviation "n.s." indicates that no significant difference was found. Among the most important takeaways from these results are that students in the online synchronous group (S) performed significantly worse on the assessment than all other groups save for Group V (asynchronous online video). The students from Group S also scored significantly lower than all other groups in terms of the average quality of their contributions to the class discussion. In terms of participation rates, however, Group S was not significantly different than the other groups, while the groups that met for in-person discussions (C and H) had significantly lower percentages of students making contributions than the asynchronous group that used video (Group V). Group H also had a lower rate of study completion than

either Group V or T. Finally, in terms of the survey of satisfaction and its sub-categories, the general trend was that the asynchronous groups typically indicated lower satisfaction than the in-person and hybrid groups (C and H, respectively). Overall, the analyses from this chapter paint a picture that indicates very different results for the different instructional strategies depending on which student outcome was being measured. The next, and final, chapter of this study will examine the implications of these findings.

Chapter 5: Discussion and Implications

Introduction and Study Synopsis

Purpose of the Study

This study set out to investigate ways that different strategies of online instruction may affect several key student outcomes. This subject is particularly topical, as annual growth in the number of students taking online courses has averaged double digit percentages since the turn of the millennium, and has easily outpaced the overall growth rate for U.S. higher education over the same timeframe (Allen & Seaman, 2014). Within this new medium of internet-based course delivery, a variety of different instructional strategies have proliferated, many with notably distinguishable characteristics. These distinctions include whether instruction is delivered in real-time or asynchronously, what media is used to deliver instruction, whether communication occurs verbally or through text, and quite a few others (Kearsley, 1997). Unfortunately, as the number of online programs has continued to grow, faculty and administrators have had very little empirical information to use in comparing and distinguishing between outcomes of these myriad strategies of online instruction. As a result, key decision makers often have little or no understanding about how their students may be affected by the decision to employ one online instructional strategy over another (Maguire, 2005; Parthasarathy & Smith, 2009). The consequence of this lack of information is that administrators and faculty are more likely to base decisions related to online instruction on guesswork, anecdotal information, or upon more strategic considerations such as costs or convenience, without knowing the potential consequences for student learning or satisfaction (McDonald, 2002).

The goal of this study was to compare several different online instructional strategies across a number of key student outcomes in order to provide evidence that might help to inform the decision-making processes of administrators and faculty who are tasked with developing online programs and courses. By conducting an experiment in which the strategy of instruction varied but other key elements of instruction (lesson, materials, instructor, timing, etc.) were kept as constant as possible, this study hoped to show how and in what capacities certain online instructional strategies might outperform others. Hence, the research question that this study sought to answer: "Does the type of online instructional strategy employed have an effect upon key student outcomes?" To satisfactorily address this question would allow educators to use the information and results of this study, in conjunction with other applicable literature and strategic considerations (i.e. cost, development time, student needs, and reusability) to make better-informed decisions about what online strategies are best suited for their courses and programs.

Prior Studies

This is not to say, of course, that there has not been a great deal of research into online education thus far, but much of this research has focused solely on comparisons between online and traditional classroom instruction. These early comparison studies do offer valuable insight into the general viability of online instruction by providing a great deal of evidence that online education can be as effective as traditional classroom instruction (Russell, 1999). Among the most notable contributors to this conclusion were several large meta-analyses which examined dozens of quasi-experiments in which a single strategy for online instruction was compared to the traditional classroom experience (Means, Toyama, Murphy, Bakia, & Jones,

2010; Sitzmann, Kraiger, Stewart, & Wisher, 2006; Jahng, Krug, & Zhang, 2007; Allen, Bourhis, Burrell, & Mabry, 2002; Byoun & Zoljargal, 2012). The fact that these meta-analyses typically showed no significant differences overall between the student outcomes from the online and face-to-face comparison groups bolster the early claims of Richard Clark (1983), who argued that students' learning outcomes are not affected by the media used to deliver the lesson.

On the other hand, many of these meta-analyses comparing online and classroom instruction share common limitations in terms of the insight that they can provide to practitioners. Because many of the studies that the meta-analyses used were quasi-experiments in which two sections of the same course were taught using online and face-to-face instruction, self-selection bias, arising from students choosing which of the sections they wished to join, may have frequently come into play (Phipps & Merisotis, 1999; Joy & Garcia, 2000; Amlie, 2003). In addition, these meta-analyses tend to take a somewhat monolithic view of online instruction by comparing in-person instruction to an over-generalized conception of online instruction that fails to take into account the diversity of online instruction strategies. Thus, it is unclear whether some online strategies may have fared significantly better or worse than others when compared to the classroom.

Moreover, other studies pertaining to online education seem to refute Clark's (1983) claim that the media used in instruction is unimportant to student learning. Among these are studies which demonstrate that different subjects can be taught more or less effectively online (Xu & Jaggars, 2013b; Al-Shalchi, 2009) or that students with different characteristics may have varying levels of success in online classes (Figlio, Rush, & Yin, 2010; Eom, Wen, & Ashill, 2006). Perhaps even more to the point are a small number of studies that compare different online

instructional strategies with each other and have found significant differences in several important outcomes (Ge, 2012; Choi & Johnson, 2005; Hrastinski, 2008a). One reason that this last group of studies is still quite small in number may be due to the difficulty and cost (Arizona Learning Systems, 1998) of implementing multiple strategies of online instruction across different sections of a course. Nevertheless, all of these studies, in the aggregate, offer strong indications that Clark's hypothesis may not hold true in all respects within today's diverse world of possible approaches to online instruction.

Investigative Approach and Analysis

This study took the view that it would be necessary to conduct a true experiment in order to better understand whether certain strategies of online instruction may differ in terms of their effects on student outcomes. Conducting the experiment outside of a pre-existing college course and randomly assigning the participants to treatment groups would make it possible to address the potential problem of self-selection bias that is present in many quasi-experimental studies of online instruction (Phipps & Merisotis, 1999). Extending this idea, it would also be beneficial to compare several different instructional strategies (including traditional classroom instruction as a control) at once across a number of key student outcomes. This would provide a much greater scope of comparison and offer the opportunity to gain useful insights into how certain instructional strategies might outperform or fall short of other approaches. This line of thinking lead to the resulting experiment: a single lesson in which the strategy of instruction varied across five treatment groups while other important elements of the student experience were held as constant as possible. All five treatment groups received instruction from the same instructor, on the same subject, using identical content, during the

same week, using a similar hour-long timeframe for instruction and discussion. Potential differences in participant characteristics were controlled for by the random assignment of participants to treatment groups.

These participants were all undergraduate students at a single large, West-Coast, public research university. The students were compensated $30 for participation in the study and were recruited through email, flyers, social media, and the campus newspaper. As part of the process for enrolling in the study, all participants filled out a demographic questionnaire (Appendix A) and verified their availability for the day and time when the study's live events were scheduled to occur. A total of 425 students enrolled in the study and were randomly assigned to treatment groups. Of these 425 original registrants, 303 completed the study.

The treatment groups to which the participants were randomly assigned represent some of the most important current strategies for online instruction. A control group (C) experienced the lesson and discussion in a classroom format; an online synchronous group (S) participated via videoconference; a hybrid group (H) watched a pre-recorded video then attended class for a discussion; an asynchronous group (V) watched a pre-recorded video then discussed in an online forum; and another asynchronous group (T) viewed the lesson as an annotated slideshow and also carried out a discussion in an online forum. The outcome measures for each group included a test of comprehension, rates of participation and quality of comments, a survey of satisfaction, and overall rates of study completion. Each of these outcomes was analyzing using ANOVA or chi-square tests with between-group comparisons for the treatment groups, regressions with variables for both the treatments and demographic

characteristics, and additional regressions to test for the presence of interaction effects between the treatments and demographic variables.

Summary of Findings

For every outcome that was measured by this study, ANOVA and chi-square consistently provided evidence, significant at a confidence level of 99%, for the existence of differences between the treatment groups. For the outcome measure of comprehension, which was evaluated using a short online quiz (Appendix B), the control group (Group C) scored the highest, while the synchronous group (Group S) received the lowest scores. The synchronous online group also scored significantly lower (at 95% confidence) on the quiz than the hybrid group (Group H) and the group which viewed their lesson as an asynchronous slideshow (Group T). In terms of engagement, the participants in the group that used asynchronous video had significantly higher rates of contributing to the class discussion than the in-person and hybrid groups (C and H). Also, the synchronous online group received significantly lower ratings for the quality of their comments than all of the other groups. The post-lesson survey of satisfaction (Appendix C) had a number of sub-sections but, in terms of overall results, the control and hybrid groups provided the highest ratings, while the two asynchronous groups were generally least positive about their experiences. Lastly, in terms of study completion, the hybrid group had a significantly lower rate of participation than either of the asynchronous groups. In light of these many observed differences, this study can answer the question of whether different online instructional strategies have varying effects upon key student outcomes for a single lesson with a confident “yes.”

Discussion of Findings

Overview of Notable Observations

The most notable finding from this study is that, in the context of a single lesson, the strategy of online instructional strategy that is employed may have a significant effect upon several important student outcomes. Moreover, there does not appear to be a single best strategy for enhancing all possible outcomes; each of the different strategies showed varying strengths and weaknesses across multiple outcomes. On the other hand, there were some strategies from the study that generally performed better or worse than others on a larger number of the measured outcomes.

The control group (C), which conducted both the lesson and discussion in a traditional classroom environment, significantly outperformed two out of the three fully-online groups on the assessment of comprehension and both of the asynchronous groups on the survey of satisfaction. The only measure for which the control group was significantly below any other groups was in the rate of participation, where it was outperformed by both asynchronous groups. The hybrid group (H) also had similarly-strong overall results. There were no significant differences between the hybrid and control groups on any outcome measures, though the hybrid group was not as far above the other groups on the assessment of comprehension, and the hybrid group was outperformed by both asynchronous groups in terms of overall rates of study completion. The synchronous group (S), on the other hand, had the lowest score on the quiz and was significantly below all of the other groups on the measure of average comment quality. The only area in which the synchronous group had any significant edge over any of the other groups was in sections of the survey of satisfaction, where its participants generally gave

higher overall ratings than the two asynchronous groups. The two asynchronous groups (V and T) were each more of a mixed bag, with strong rates of participation and comment quality but the lowest overall ratings on student satisfaction. Interestingly, despite using very different media to deliver the lesson, there were no significant differences between the two asynchronous groups on most outcome measures, save satisfaction with the instructor, in which the group that used video (V) rated their satisfaction as significantly higher than the group that used only slides and text (T).

Clearly, these findings offer evidence against Clark's (1983) argument that the medium of instruction does not affect students' learning outcomes. The results from this study also do not conform to Russell's (1999) theory of *no significant difference* for student outcomes when the independent variable is the strategy of instruction. On the other hand, it is not clear that the results from this survey contradict the majority of other studies examining online education. After all, the large meta-analyses such as Means, Toyama, Murphy, Bakia, and Jones (2010), include both studies in which there were no significant differences between instructional strategies and studies in which there were differences in both directions. In addition, since these meta-analyses generally examine only comparisons between face-to-face and single instances of online instruction, broadly defined, it is difficult to compare their results with this study's comparison of several, specifically-defined instructional strategies. It is, in fact, difficult to compare the results of this study with any study that did not differentiate between strategies for online instruction. Later parts of this chapter will compare how the results of this study match with the small number of studies that have made comparisons that distinguish between online instructional strategies.

Quiz Results

The post-lesson online quiz (Appendix B) that was administered to all of the treatment groups in this study was intended to gauge the participants' comprehension of the lesson and their ability to offer proof of learning. According to the quiz results, the group that participated via online videoconference (Group S) learned the least about the subject matter of the lesson. Group S's average score was the lowest, and was significantly below all other groups except the asynchronous group that watched the lesson via pre-recorded video (Group V). Although the synchronous online group (S) also had the lowest average quiz score in the pilot study, this result is still somewhat surprising in many ways, not the least of which being that many educators would consider synchronous online instruction to be the online education strategy that is most similar to a traditional classroom environment (Lieblein, 2000). Also, while participants in the other online groups (H, V, and T) could easily have cut corners by viewing only a small portion of the instructional material provided to them, participants in the synchronous online videoconference were, with the exception of a few late (<10 min.) arrivals, logged in for the full hour of lecture and discussion.

This is not to say, however, that the students in Group S were necessarily paying attention for the entire webinar. Aside from comments in the chat window, there was simply no way to know whether participants in Group S were browsing other sites were even present at the computer while the videoconference lesson was ongoing. This possibility, of course, would also be the case in most learning environments in which synchronous instruction was employed for a similarly-sized group of students. Another issue to note concerning the synchronous online group's performance was the presence of minor technical problems,

including two brief losses of the video image when the instructor's screensaver activated and audio configuration issues that affected some participants at the beginning of the lesson. While these glitches did not have an impact in terms of timing, it may be that the interruptions subtly altered the flow of the lesson or upset the students' attention. Of course, technical problems are an inherent risk of using videoconferencing technology and minor glitches are very common in synchronous online learning environments (Wang & Hsu, 2008). Taken from that perspective, the somewhat-buggy experience of the participants in Group S may have been a fairly realistic representation of the instructional strategy being evaluated.

It is worth noting that this study does not provide the first observed evidence that asynchronous and hybrid instruction may outperform synchronous instruction in terms of learning outcomes. Bernard et al (2004), also noted this result in their meta-analysis of comparison studies between distance education and face-to-face instruction. In their examination of 232 of these comparison studies, they found that asynchronous methods of instruction tended to outperform the classroom, while synchronous studies tended to underperform. Representing the other side of this issue are researchers, such as Skylar (2009), who found no significant differences in learning outcomes between asynchronous and synchronous treatment groups. Interestingly, however, this study's literature review found no other studies that showed significantly positive results for synchronous online instruction when compared to asynchronous in terms of quantifiable learning outcomes.

One more important observation from the quiz results was that the classroom control group (C) significantly outperformed the online asynchronous group which viewed the lesson via pre-recorded video (Group V), but not the online asynchronous group which viewed the

lesson through annotated slides (Group T). While there was no significant difference between the two asynchronous groups' quiz scores, this is an interesting observation in itself, since it seems to be contrary to Choi and Johnson's (2005) observation that students are more motivated and attentive in video-based lessons as compared to text-based. The fact that the classroom control group significantly outperformed even one of the asynchronous groups also raises an interesting question as to why this would be the case.

As previously mentioned, the instructor who took part in this study is a veteran lecturer who is highly experienced in delivering the presentation that was used by the study. He used the same presentation slides for all treatments and, because of his familiarity with the subject, was able to deliver a very similar performance in his classroom lecture and pre-recorded video. The production quality of the video, while not up to the standards of modern television, was nonetheless perfectly clear and easy to understand. One possible explanation for the observed differences in comprehension between the asynchronous video-based group and the classroom control could be that the questions which were asked by the students in person added to the control group's understanding of the material. It could also have been that the students in the control group found it easier to pay attention in person compared to the students who watched the lesson on their computers. A third possibility is simply that some of the students in the asynchronous group did not watch the entire lesson or attempted to multitask while the video was playing. To bolster this last theory, it is worth noting that the hybrid group (H) watched the exact same video as the asynchronous group, but the hybrid group's results showed no significant difference from the control. Despite having access to the exact same video, the difference in the scores between the hybrid Group H and asynchronous Group V could be

explained by the fact that the hybrid participants knew that they would eventually be sitting in front of the instructor in a classroom and might have to answer questions about the lesson.

The idea that differences in the participants' quiz scores may be largely attributed to varying degrees of accountability between the instructional strategies may warrant further exploration and will be discussed later in this chapter. For now, it is interesting to speculate about how results may have been different had the two asynchronous groups viewed their respective lessons on computers, but in a controlled classroom or lab environment with an instructor present, as opposed to in the settings of their choice. Also, might the inclusion of mandatory "check-in" questions sprinkled throughout the asynchronous lessons have helped to keep the participants in Groups V and T on track and accountable? In a similar vein, what would have happened if participants in the online synchronous group had been required to turn on their computers' webcams during the webinar? It is not unreasonable to suppose that participants' engagement and attention might have increased if they knew they could be observed.

Engagement

This study evaluated participants' engagement with the lesson using two measures: rates of participation and average comment quality. The former measure showed a marked difference between the two asynchronous online groups and the two groups which met in person (control and hybrid). In particular, the asynchronous group that used video (V) had a significantly higher (at alpha=0.05) percentage of participants who submitted comments or questions than the control group (C) or the hybrid group (H). These results are not surprising for a number of reasons, perhaps the most apparent being that the asynchronous groups had a

full week during which they were able to submit comments or questions via discussion forums. The other groups, by contrast, only had a limited window in which to make comments: one hour in the case of the control and synchronous groups (C and S) and just twenty minutes in the case of the hybrid group (H).

Another common sense reason that might help to explain the higher rates of participation from the asynchronous groups is that they were able to make simultaneous comments. After all, participants from Groups V and T could have been online at the same time posting very similar comments to one another without being aware of each other until they saw the other's posts. Basic etiquette, on the other hand, dictated that participants in the groups which met in the classroom (C and H) not speak over each other or answer questions all at once. Interestingly, the same etiquette was not a constraint for the synchronous online group (S). Because synchronous participants were unaware when others were typing comments, it was not uncommon for multiple comments to be posted nearly simultaneously. This may explain why Group S, which did have a narrow window of time in which to make comments, but was not constrained in terms of simultaneous participation, was very near the study's average overall rate of participation and was not significantly different from any other groups on this measure.

One additional factor that may have influenced the participants in the asynchronous groups to participate at a higher rate was the way that their discussion forums were built into the lesson websites for Groups V and T. Though students in all groups were presented with the opportunity to ask questions and make comments on discussion topics without participation being described as mandatory, the fact that the discussion forums were visibly part of the

lesson websites for Groups V and T may have made the discussions seem like a more formal or required process. It is also worth noting that participants in the asynchronous groups could see a count of how many new posts were being made in the discussion forums, so it is possible that the students in these groups may have felt some pressure to participate or were encouraged to participate based on the number of other posts that were being made.

Due to the reasons cited above, the higher rates of participation from the asynchronous online groups compared to the control group were to be expected. It was more unexpected to observe that the online synchronous group (S) scored significantly below (at alpha=0.05) all other groups in terms of comment quality. This may be partly because, as previously noted, comments from the participants in Group S who participated via videoconference were quite a bit shorter, on average, than any of the other groups. But this, in turn, begs the question: why were the comments made by the students in the webinar (Group S) so much shorter?

One possible reason that the comments from the webinar were shorter and of generally poorer quality is that there was considerably less structure and formality in the video chat as compared to a live classroom or discussion forum. Because participants in the videoconference could not tell when their classmates were typing, they may not have felt constrained to wait for their turn to "speak." Also, when the instructor asked the class a question, he generally responded to the first response to appear. This may have created pressure to send in an answer first, which might have discouraged thoughtful review or elaboration from the participants. Also, although participants were identified in the webinar by name, the videoconference format may have contributed to a feeling of anonymity among the synchronous online participants for the simple fact that their faces could not be seen. Unlike in

the discussion forums used by Groups V and T, the posts from participants in Group S may have seemed less permanent because of the rapid, running flow of the comments. In other words, participants in the webinar may have felt less embarrassed about providing a bad, incorrect, or partial answer because their comments were not prominently displayed for more than a few seconds. They may even have felt a there to be a different, more casual, standard for what might be considered a rude response in the chat environment. For instance, responding to an instructor's question with a single-word answer in a classroom environment might be considered by some students to be somewhat rude or dismissive whereas in a chat session it may appear to match the pace and flow of the discourse.

Another possible reason that the online synchronous group (S) scored significantly lower than the other groups in terms of the quality of their comments may be that the pace of the discussion was faster in the videoconference. Considering that there was no significant difference in participation rates between Group S and the asynchronous online groups, despite the fact that Groups V and T had a full week to comment while the participants in Group S only had an hour, it seems possible that the pace of the discussion in Group S may have encouraged rapid, lower-quality responses. Of course, the other possibility is that the students in the videoconference were simply not as engaged by the lesson and hence did not put in as much effort into their comments as the other groups.

This study's approach to comparing the quality of class contributions across instructional formats may be unique to the field of online education research. There have been other studies, however, that have examined other outcome measures with the format of the discussion (synchronous vs. asynchronous) as the independent variable. Cleveland-Innes and

Ally (2013), for example, found no significant difference between synchronous and asynchronous online discussion formats in assessments of comprehension. Hrastinski (2008a), in his investigation into the differences between synchronous and asynchronous online communication found synchronous communication to be more social and interactive when compared to the more thoughtful and complex comments made in asynchronous formats. While this study obviously did not investigate these same attributes of discussions as Hrastinski, the results of this study do not seem to be out of line with Hrastinski's conclusions. Taken together, these studies suggest that different mediums of online communication have different norms that dictate different experiences for their users. The implications of some of these experiential differences will be discussed later in the chapter.

Satisfaction

The results of the post-lesson survey that was administered to all participants (Appendix C), indicated wide differences in students' satisfaction with their experiences in the study, depending on the strategy of instruction to which they were exposed. Although the survey measured several specific areas of satisfaction, there were some general trends that applied to most questions in the survey. For instance, the students in the classroom control and hybrid groups (C and H) tended to rate their satisfaction as highest overall, while the students who were in the groups which received their instruction asynchronously online (V and T) tended to rate their satisfaction the lowest. The online synchronous group (S) was typically somewhere in the middle on most ratings of satisfaction; for the overall survey average their ratings were only significantly different from the group which received their instruction through asynchronous online slides (T). In fact, all other treatment groups, save the other asynchronous group (V)

were higher than Group T in terms of their overall average ratings on the satisfaction survey. The most obvious conclusion from these results may be that students prefer direct contact with each other and/or the instructor during the course of their lessons. Hence, the greater their transactional distance (Moore, 1993) in the lesson, the lower the students' average ratings of satisfaction. While this trend varies somewhat by degrees according to the area of satisfaction being measured, it generally holds true across all sub-sections of the survey.

The first of these survey subsections asked participants to evaluate the subject matter of their lessons. Not surprisingly, since the subject was the same across instructional strategies, there was not as much difference between the treatment groups as there was for the other survey subsections. Nevertheless, the hybrid group (H) did submit ratings for the subject matter that were significantly higher than both asynchronous online treatment groups (V and T). This is especially interesting because the only difference between the instructional strategy for the hybrid group (H) and the asynchronous online group that viewed the lecture as a video (Group V) was the medium in which the discussion questions were presented. Both groups watched the identical video for the lecture portion of their lesson, which was the primary way in which the subject was presented. However, the hybrid group's in-person meeting for discussion seems to have influenced their perceptions of the lesson's subject matter (particularly how interesting and useful they felt it was), or at least their ratings of the same. This may indicate that differences between these groups' ratings for the subject matter may have been, more than anything else, a referendum on their overall satisfaction with their experiences in the lesson.

While the fact that there were significant differences between the groups' ratings of the subject was somewhat surprising, by contrast, it was absolutely expected that there would be major differences in terms of the groups' ratings of the instructor. Even though the instructor, like the subject matter, was the same across all treatment groups, the context of each group's exposure to him was so different that divergent levels of satisfaction were anticipated. These expectations were borne out with strongly-significant differences between the treatment groups' ratings of the instructor, particularly with the control and hybrid groups (C and H) giving him much higher ratings than the two asynchronous groups (V and T). Instructor ratings were so different, in fact, that the control group's ratings were found to be significantly higher than the online synchronous group (S) which, in turn, gave significantly higher ratings than the online asynchronous group that used annotated slides (T). Since Group T never saw the instructor, even as a video image, that group's significantly lower instructor ratings compared to all of the other groups was also no great surprise, even though the content of the instructor's lesson was largely the same. All of these results may indicate that students value seeing their instructor in any format, and that fewer barriers to direct communication between the students and the instructor are also preferred. This, in turn, raises another question: would the differences in satisfaction across groups have been less-pronounced if the instructor had been less charismatic– or vice versa? In other words, if the classroom students had valued their time with the professor less, would there have been a smaller difference between group satisfaction ratings? Conversely, if the instructor's ratings with the in-person group had been higher, would the between-group differences have been greater?

Like the differences between the groups' instructor ratings, it was also anticipated that there would be significant differences in how the different groups rated the format of their lessons. Interestingly, however, the differences in format ratings were actually less-pronounced than the differences in instructor ratings, despite the actual differences in format between groups. The only significant difference for format ratings was that the control group (C) awarded higher scores for satisfaction with format than both asynchronous groups (V and T). This, in particular, was not a surprising result. Since all study participants were students who opted to complete their undergraduate studies in a traditional, classroom-oriented format, it was to be expected that they would prefer this strategy of instruction to others. It is, perhaps, more interesting that there were no significant differences between any of the other treatment groups. This suggests that students who are accustomed to classroom instruction may not have strong outward preferences for any given type of online instructional strategy. On the other hand, the fact that other measures of satisfaction do show significant differences indicates that, even if these students do not have explicit preferences as to the format of their online instruction, the attributes of the particular format that they are exposed to do influence their satisfaction in other ways.

One of the ways in which the format of instruction influences a specific aspect of the students' experience in the lesson is with regards to interactions. In terms of the students' satisfaction with their lessons' interactivity, there was a significant split between the groups that conducted their lessons with real-time interactions (C, S, and H) versus those that interacted asynchronously (V and T). All of the groups that experienced real-time interaction rated their satisfaction with their lessons' interactions significantly higher than the groups that

interacted asynchronously. There were no significant differences between Groups C, S, and H or between Groups V and T. These results strongly suggest that real-time discussions, even if conducted online, are more popular with students than asynchronous discussions. While this result is, on its own, not entirely surprising, it is interesting to contrast the low satisfaction ratings on interactions from the groups which communicated using asynchronous forums (V and T) with the observation that these same groups had the highest rates of participation in discussions and the highest average ratings for comment quality.

Finally, in the last section of the survey, which related to participants' general satisfaction with the lesson and their willingness to experience more of the same, it was not surprising to see that the broader trends in satisfaction from the overall survey were again borne out. In this last category, the control group (C) again had the highest average ratings of satisfaction, which were significantly higher than both asynchronous groups (V and T). As previously noted with regard to the format of the lesson, this result was expected, since all of the participants who took part in the study were engaged in a traditional, campus-based, undergraduate experience and one of the survey questions in this section asked about the participants' interest in signing up for another course of the same type. Unlike with the results of the questions pertaining to format, however, there was one significant difference between the online treatment groups. Namely, the hybrid group (H) rated their satisfaction with the learning experience significantly higher than the asynchronous group which viewed their lesson using annotated slides (Group T). More than anything else, this highlights the fact that, among the online treatment groups, the hybrid group had the highest ratings of satisfaction on every measure, while one of the asynchronous groups always had the lowest.

These results are in-line with some, but not all of the research on satisfaction in online courses. Several studies, such as Benoit et al (2006), Cole et al (2012), and Johnson, Aragon, and Shaik (2000) similarly found that students in fully-online courses tend to be less satisfied with their experiences than students in the classroom. On the other hand, Shelley, Swartz, and Cole (2007) as well as Allen et al (2002) both found no significant differences in terms of students' satisfaction in online and face-to-face courses, while Eom, Wen, and Ashill (2006) found online students to be more satisfied. Unfortunately, in the above-referenced studies, it is not possible to determine the online instructional strategies that were being evaluated against the classroom experience, except in the case of Shelley, Swartz, and Cole (Shelley, Swartz, & Cole, 2007), which states that the online students used both synchronous and asynchronous methods of communication. Given the significant differences in satisfaction displayed by the treatment groups in this study, there appears to be ample evidence that, in terms of student satisfaction, the strategy of online instruction is an important factor that should not be ignored when comparing online and face-to-face instruction. Finally, it also worth pointing out that the results of this study appear to provide support to Moore's (1993) Theory of Transactional Distance, at least in terms of its relationship to student satisfaction. In line with Moore's theory, the asynchronous groups in this study, which would be associated with having the greatest transactional distance, generally rated their satisfaction the lowest.

Study Completion

The final student outcome examined by this study was the rate at which the registrants assigned to each of the treatment groups completed the study. Since participating in the group discussions was not required, study completion was effectively a measure of the rate at which

students in each of the treatment groups both engaged with the lesson and completed the post-lesson survey and quiz. Comparing this rate across treatment groups found the highest rates of completion in the two asynchronous groups (V and T) and the lowest rate of completion in the hybrid group (H). The differences between the asynchronous and hybrid groups were the only significant group differences for this outcome. It can be speculated that the asynchronous groups had the highest rates of completion because they benefitted from the convenience of having an entire week in which to view their lessons. This flexibility may have allowed them to work around most potential scheduling conflicts. The hybrid group, on the other hand, may have had the lowest rate of completion due to their dual responsibilities to watch the lesson via video and attend the group discussion in person.

The higher lesson completion rates from the two online asynchronous groups in this study seems to contradict a good deal of prior research that indicates that persistence rates are lower in online courses (Brady, 2001; Carnoy, Rabling, Castano-Munoz, Montoliu, & Sancho-Vinuesa, 2012; Diaz, 2002; Gleason, 2004). There is, however, a critically-important distinction between this study's results and the results of prior studies which have dealt with online student retention—namely, the fact that this study only examined a single lesson as opposed to an entire course or program. Hence, it is probably not advisable to infer too much into this study's completion rates as far as they might relate to student retention over an entire course or program.

Essentially, due to the single-lesson nature of this study, the online asynchronous groups may have had certain advantages compared to the other groups with regards to completion of the study that probably would not translate to a longer-term curriculum. To wit,

the duration and amount of activities in the study's single lesson was not extensive enough to expect students to drop out due to dissatisfaction with their learning experiences and, due to the previously-cited differences in rates of satisfaction, it may be presumed that a longer, more extensive asynchronous program would have had more of a negative effect on the relatively-dissatisfied asynchronous groups. In addition, the effects of procrastination, which have been shown to disproportionally reduce student performance in asynchronous online courses (Elvers, Polzella, & Graetz, 2003) also cannot have been as much of a factor for the single lesson. On the other side of this equation, it is reasonable to assume that the live sessions in this study might have been better-attended if they had been part of the students' regular weekly schedules. The fact that this lesson was a one-off event may have meant that it was more susceptible to conflicts in the students' calendars. All of this having been said, the convenience factor for asynchronous instruction cannot be ignored. At least for single lessons, the results of this study suggest that offering those lessons in an asynchronous online format may be the best strategy for garnering the highest rates of completion.

Interaction Effects

With certain exceptions noted in Chapter 4, the results of this study generally did not provide evidence of interaction effects between strategies of instruction and the participants' demographic information. In other words, there was little to suggest that the demographic variables (in most cases) enhanced or mitigated the effects of the strategy of instruction on student outcomes, and vice versa. This lack of significant interactions is particularly notable in the cases of GPA, gender, and race—each of which have been shown by prior studies to interact with online instruction to affect student learning outcomes (Xu & Jaggars, 2013b; Figlio,

Rush, & Yin, 2010). More specifically, these prior studies showed that students with low GPAs, males, and under-represented minorities fared disproportionally worse in online classes compared to other students. This study was unable to reproduce these interaction results. One possible explanation for this could, again, have to do with the inherent differences between one lesson and a full course. Other possibilities include self-selection biases in the prior studies, both of which used quasi-experimental designs, or the differences in the populations from which each study's participants were drawn.

Implications for Practice

One of the most apparent implications of this study's results is that it is not an easy task for standard online instructional strategies to exceed, or even to equal, the success of classroom instruction in terms of the effect on many important student outcomes. Given the tried-and-true nature of traditional classroom instruction, as well as the massive infrastructure within the U.S. higher education system to support it, traditional classroom instruction may be the simplest and most-direct way for many schools to optimize most aspects of student success. It is therefore quite possible that some of the recent prognosticators who have predicted the imminent demise of a large portion of campus-based programs at the hands of online competitors (Christensen & Eyring, 2011; Byrne, 2014; Cuban M. , 2012) may have jumped the gun as long as student educational outcomes and satisfaction continue to be major factors in the success of most post-secondary programs. On the other hand, it is impossible to ignore the advantages in convenience (Hrastinski, 2008b) and scalability (Laws, Howell, & Lindsay, 2003) that some strategies of online instruction have over traditional instruction—as

highlighted by the high rates of participation and study completion demonstrated by this study's asynchronous online treatment groups.

The results of this study also show that hybrid or blended strategies for instruction can also have very strong results. Because there were no significant differences between any of the hybrid treatment group's results and the results from the classroom control group, it can be surmised that the brief in-person contact that the hybrid group experienced went a long way towards mitigating the comparative dissatisfaction experienced by the asynchronous online groups. By incorporating some online elements, hybrid instruction may also partially attain some of the advantages in convenience, scalability, and long-term cost savings (Bowen, Chingos, Lack, & Nygren, 2012; Benoit, Benoit, Milyo, & Hansen, 2006; Vilaseca & Castillo, 2008) of online instruction. Finding the "ideal" mix of various online and in-person instructional techniques to match their students' needs may be a critical challenge facing educators who design hybrid programs of instruction in the future. Regardless, it seems clear that some amount of direct, in-person contact may be important for optimizing student satisfaction.

With regards to this last point, it is important to distinguish between the outcomes associated with in-person contact versus synchronous online communication. Although both strategies involve live interaction, the results of this study suggest that in-person and hybrid instruction have superior results to videoconferencing on several key student outcome measures—most notably on the assessment of comprehension. Perhaps even more significantly, with the exception of some measures of satisfaction, the synchronous treatment group in this study failed to outperform the two asynchronous groups on most of the measured student outcomes. In fact, the online synchronous method of instruction resulted in both the

lowest average quiz scores and the lowest average scores for comment quality across all instructional strategies evaluated by this study.

This finding could have important implications for the design of online programs and courses. Although synchronous online instruction may seem less foreign to instructors who are used to teaching in the classroom (Lieblein, 2000), it does not seem to be superior to asynchronous online instructional strategies with regards to most important student outcomes, and may be decidedly inferior in others. In addition, it sacrifices a great deal to asynchronous strategies in terms of convenience, reusability, scheduling flexibility, demands on faculty time, and the ability to scale to accommodate large numbers of students (Christensen & Eyring, 2011; Aslanian & Clinefelter, 2013; Laws, Howell, & Lindsay, 2003; Skiba, 2012). If all that synchronous online instruction can offer over asynchronous strategies are some measures of satisfaction and an easier set-up process (Laws, Howell, & Lindsay, 2003), then in many cases it may not make sense to base an entire course or program on strictly synchronous online instruction. Instead, for most online teaching environments, it may be advisable to utilize synchronous tools primary for interactive communication and to rely on asynchronous tools to form the basis of instruction and course delivery.

With regards to asynchronous online instructional strategies, another interesting takeaway from the study is that there were not many significant differences between the outcomes attained by the two asynchronous groups, except in terms of their satisfaction with the instructor. In other words, in most respects, it did not seem to make a major difference whether the students viewed their lesson as a pre-recorded video of the instructor or as an annotated slideshow. Of course, both methods were low on ratings of student satisfaction

compared to the other groups, but the fact that the slides performed as well as the video on most outcome measures might be surprising to some practitioners. This seems to indicate support for Clark's (1983) argument that, so long as the information is clearly presented, the medium may not have a significant effect on learning or other student outcomes.

On the other hand, it is worth pointing out that, as video production costs continue to fall (Fox, 2010), there may not be significant cost savings associated with using static asynchronous content over video. All the same, annotated slides and other non-video alternatives may be preferable in certain situations or geographic locations where high bandwidth internet connections are not available and video streaming is not feasible. In addition, there may be advantages associated with employing a media mix for instruction instead of simply relying on a single medium. For instance, it is reasonable to consider whether the lower instructor-satisfaction ratings from Group T might have been considerably higher if the lesson had begun with a short introductory video (with the instructor featured prominently) to go along with the annotated slides. Also, might a hybrid strategy that employed slides instead of video for instruction, but still met in person for discussion, have produced similar results as compared to Group H? While this study cannot answer these questions, the implication that video is not necessarily a superior medium for asynchronous instruction opens up an array of interesting approaches to online instruction using different types and combinations of asynchronous tools and media.

One such tool, commonly used in asynchronous online classes, is the discussion forum, which this study showed to be effective in fostering both high levels of participation in class discussions and thoughtfully conceived student comments. The results of this study suggest

that online forums may be a good way to encourage large numbers of students to engage with the content of the lesson. The detailed and considered comments collected from the students in this study's asynchronous treatment groups also suggest that students may be more motivated to make high-quality comments (or at least to avoid making bad comments) in online forums. On the other hand, the results from this study also indicate that students are less satisfied with this form of interaction. Based on these observations, one promising approach may be to use asynchronous discussion forums as one means to engage students in traditional and hybrid classes. Using this approach, the online forums could be employed as a lead-in to classroom discussions in order to involve larger numbers of students in the conversation and to encourage broader consideration of key instructional concepts. Meanwhile, other forms of in-person interaction could help to address the lower satisfaction rates associated with asynchronous online communication. This is just one possible way in which online discussion forums could be used; the main takeaway from this study's results is that these asynchronous forums are potentially valuable and versatile tools that may be effectively employed to foster broader engagement in a variety of contexts.

This example highlights what is probably the biggest overall implication for practice from the results of this study; one that is based on the observation that different online instructional strategies have various strengths and weaknesses according to the student outcomes being measured. More to the point, there does not seem to be one "best" strategy for online instruction, but rather a number of more favorable ways to tailor programs depending on the circumstances and goals of instruction. This observation leads to the conclusion that different media mixes and combinations of operational approaches may be best

able to exploit the relative strengths of different instructional strategies. Depending on the needs and goals of the program, there may be circumstances when one of the five instructional strategies evaluated by this study may be the best approach to use. On the other hand, there may be situations when it makes more sense to combine elements from different strategies in order to exploit their various strengths. Examples of these combined approaches could include fully-online classes that use pre-recorded videos to deliver lectures and videoconferencing for interactive sessions, classes taught on campus that incorporate online discussion forums to enhance engagement, or hybrid classes that also utilize videoconferencing to hold online office hours. Of course, these are not new ideas, but the results of this study provide evidence that not all combinations of instructional techniques are created equal. The key, it seems, is in picking the right tools and techniques to match both the strategic aims of the lesson and the student outcomes that are most highly valued.

Limitations

There are several limitations to the practical conclusions that can be drawn from the results of this study. The most notable limitation is that this study is based on only a single lesson—as opposed to an entire class. While this approach allowed the analysis to focus more narrowly on specific learning outcomes for the lesson, it is unclear which, if any, conclusions drawn from the study would hold up over an entire 10- to 15-week-long course. It is possible that student outcomes for certain instructional strategies could change over a longer timeframe. If that is the case, the differences between treatments observed by this study could either be diminished or enhanced over an entire quarter or semester of study.

For example, it would reasonable to assume that the relatively higher completion rates by the asynchronous groups might not hold up over a longer period of time. This supposition stems from both the literature, which typically shows lower persistence rates in asynchronous online courses (Brady, 2001; Carnoy, Rabling, Castano-Munoz, Montoliu, & Sancho-Vinuesa, 2012; Gleason, 2004), and from the assumption that it may be easier for students to stay on track in courses that have regularly scheduled meeting times each week. On the other hand, when employed over a longer period, the static nature of asynchronous content may positively affect comprehension by allowing students who have access to refer back to earlier content when studying for tests based on multiple lessons. In terms of levels of satisfaction, it is reasonable to assume that the differences in satisfaction between strategies might be magnified when polled over several lessons. All the same, regardless of whether there are differences between single lesson and full course student outcomes, this study's results do provide useful insights into outcomes for different instructional strategies on a small scale, which may then be applied, at a minimum, to the design of individual lessons within a course.

Another important limitation of this study is that its results may not necessarily apply to other subjects that could be taught using the same strategies of instruction. This study evaluated the outcomes from a lesson on the subject of law and international treaties. While it seems probable that these results would be more or less replicable for lessons on a variety of social science and professional subjects, it is an open question as to whether similar results would be obtained from a lesson on applied science, for instance. Other variables having to do with aspects of instruction are simply impossible to generalize for every possible lesson. Instructor teaching style, for example, may also interact with the strategy of instruction to

produce results different than those observed by this study. It is not unreasonable to think that one instructor might be more comfortable or would find it easier to teach using certain instructional strategies over others. Unfortunately, every possible distinct lesson characteristic simply cannot all be tested in a single experiment. Additional experiments are necessary to gain a better understanding of whether and to what extent these variables may interact with different instructional strategies to affect student outcomes.

One such critical variable related to the generalizability of this study's findings is the extent to which the students who took part in this study were representative of the entire population of U.S. post-secondary students. Simply stated, this study's participants were not representative of the majority of U.S. college students in a number of important respects. First, the participants were recruited from an institution that is both highly selective and, as a public university, draws the majority of its students from a single Western state. Additionally, the study participants had a number of important distinctions even from the overall student body of the institution where they were enrolled. As noted in Chapter 4, nearly three quarters of the study's participants were female, approximately three fifths were Asian, and four fifths were age nineteen or younger. Without any need for analysis, these characteristics are obviously significantly different from the average demographics of the entire U.S. undergraduate student body. On the other hand, it is unclear whether student demographic variables have any impact on the ways that instructional strategies affect student outcomes. The general lack of significant interaction effects in this study's results suggests that they may not. Regardless, maintaining conservative expectations regarding the replicability of these results when dealing with notably different student characteristics is advisable.

Another limitation to the study's generalizability that also relates to the characteristics of the lesson concerns the instruments used to measure certain student outcomes. The quiz and the discussion prompts utilized by this study were both adapted from a course that the study's instructor regularly teaches, though there are many possible formats that these instruments might reasonably have taken. Had the format of either the quiz or the discussion prompts been different, it is possible that the effects of the instructional strategies on the comprehension and engagement outcomes might have been notably different. For instance, the discussion prompts provided to all treatment groups in this study focused primarily on the application of important points from the lesson to an example case. If, on the other hand, the prompts had been more open-ended with no clear answer, it is worth considering whether participation rates and comment quality outcomes might have been relatively different between instructional strategies. Likewise, the quiz format, which consisted entirely of multiple choice and short answer questions, begs the question whether participants who were exposed to certain instructional strategies might have performed better or worse if the assessment had instead been in the form of an essay. It is also worth noting a potential concern related to the study's quiz instrument: that female participants scored significantly lower on average than male participants across all instructional strategies. This could be either an indication of a flaw in the design of the quiz instrument or further evidence of a lack of representativeness among the study participants. Finally, it should be pointed out that, because this study was conducted as a stand-alone lesson that had no effect on participating students' academic statuses, it is unclear whether grades, credits, and other potentially motivating factors might alter the results if these instructional strategies were to be employed in standard credit-bearing courses.

One final limitation of this study relates to which online instructional strategies were evaluated and which were left out. Unfortunately, it was not feasible for this study to compare all possible strategies for online instruction. Some of the most notable instructional strategies that were not evaluated in this study include asynchronous instruction based on adaptive assessments, fully-online instruction that blends synchronous and asynchronous elements, and interactive "game-based" simulations. Furthermore, there are many combinations of instructional elements that can be combined to form different instructional strategies. Even among the strategies which were evaluated, it is difficult to separate out the effects of different elements and techniques, such as the use of pre-recorded video for instruction coupled with an asynchronous forum for discussion. As previously noted, this study sought to evaluate the most relevant instructional strategies for modern higher education, consisting of the most commonly-used combinations of instructional elements. Nevertheless, it is clear that there are many more possible avenues to investigate.

Potential for Future Research

Because strategies for online instruction are rapidly evolving, it is important to continue to evaluate new instructional strategies as they arise. In these appraisals, it will be necessary to compare new strategies of instruction, not just with traditional classroom instruction, but also with other common strategies for the delivery of online education. Even beyond fully-developed online instructional strategies, it will also be beneficial to investigate new forms of media, instructional elements, and even software. For instance, despite their widespread adoption in higher education, the question of whether different learning management systems may have varied effects upon student outcomes is still very much unknown (Coates, James, &

Baldwin, 2005). Different combinations of media and instructional elements, including alternate models of hybrid instruction, are also worth investigating. Regardless of what is being evaluated, however, it will be important to test with randomly-assigned treatment groups in order to prevent the possibility of self-selection bias, which may have the potential to bias outcomes for courses taught using new technologies (Amlie, 2003).

In order to build upon the results of this study, future studies may wish to examine the effects of instructional strategies on student outcomes over the course of multiple lessons. Extending an experiment, such as the one carried out by this study, over multiple lessons could provide important insight into how the outcomes associated with different instructional strategies may differ from a single lesson to several lessons. Again, maintaining random group assignments will be key to obtaining reliable results—but this may present critical challenges to this type of investigation. Nevertheless, a study involving the random assignment of enrolled college students to different sections of a credit-bearing course, each utilizing a different instructional strategy, would be ideal since this type of experiment would factor in real motivations and incentives for regular college courses that were not accounted for by this study. If, however, it is not feasible to randomly assign college students to different sections of a credit-bearing course (particularly when the method of instruction varies), then it would still be useful to extend the experiment over at least three or four lessons to gauge whether an extended timeframe causes differences in outcomes across instructional strategies to increase or diminish.

To further extend this study's results, and to gain a better understanding of what may cause differences in outcomes for various instructional strategies, future studies could attempt

to isolate the operational designs of the strategies from the media that they use. More to the point, whenever online instructional strategies are evaluated, there is some question as to what extent the outcomes are affected by the manner in which students are led to interact with the instructional media versus the effects of the media itself. For example, did students in this study's control group (C) perform significantly better on the quiz than the group that watched the lesson as an asynchronous video (V) because it was easier for the students in the classroom to follow and comprehend an in-person lecture versus a pre-recorded video? Or was the difference due to the fact that the students in the classroom were a relatively captive audience while the students at home could more easily have been multi-tasking or even neglected to watch the lecture in its entirety?

One possible way to separate these two factors would be to hold the operational design of the lesson constant while varying the media. In the above example, for instance, adding another treatment group that would watch the same video as the asynchronous group in a supervised environment with the instructor present, such as a computer lab, would more closely mimic a classroom's operations. If the outcomes from this group were found to be significantly different from the outcomes from the classroom group, then it could be an indication that the media itself made the difference. Another option would be to hold the media constant and vary the operational aspects of the lesson. Continuing with the example of pre-recorded video, this could be accomplished by having one group watch a video in class, another group watch a streamed video at a specific time from whatever location they chose, and a third group download the video to watch at their convenience. Significant differences in

outcomes between these groups would provide evidence that the operational aspects of the media delivery caused the differences.

Other areas for future inquiry include evaluating online instructional strategies using different types of lessons. For instance, it would be interesting to conduct a similar experiment to the one carried out by this study but with different subject matter to see if it would yield similar results. Other types of subjects that might be evaluated include languages, technical or how-to lessons, and highly quantitative material. Another angle to investigate would involve varying the types of instruments used to evaluate the student outcomes. It is possible that students who have been exposed to different instructional strategies may perform at higher or lower levels depending on the methods of assessment used. Some possibilities in terms of different assessment methods include essay-based questions, performance of a task, or the use of a follow-up exam at a later date to test long-term retention of the concepts. Varying the size of the treatment groups might also yield interesting results.

One final area for potentially promising future research may involve conducting similar studies using different student groups as participants. Although this study found no interaction effects between demographic characteristics and strategies of instruction, the somewhat homogenous nature of the participants who took part in this study made it impossible to evaluate all possible demographic variables. It would be valuable, for instance, to include subjects from different sectors of the higher education system, including community college and graduate students. Another possibility would be to conduct an experiment to specifically test for interaction effects between instructional strategies and demographic characteristics that have been previously identified by the literature using quasi-experiments. Some of these

characteristics include race, age, GPA, gender, and learning styles (Xu & Jaggars, 2013b; Figlio, Rush, & Yin, 2010; Drago & Wagner, 2004; Eom, Wen, & Ashill, 2006). Using a well-designed experiment, it should be possible to determine whether the interaction effects observed by these studies were caused by self-selection or by true differences in how each group responds to various methods of instruction.

Conclusion

There can be no doubt that, from a research perspective, online education is a moving target. New strategies of online instruction are continually arising with the advent of innovative technologies and from different configurations of pre-existing media and instructional techniques. The results of this study indicate that educators should not assume that the effects of new instructional strategies will necessarily be the same across student outcomes. Instead, new techniques should be evaluated versus other instructional strategies to assess their respective strengths and weaknesses, and to determine in what context or combination each can be most effectively utilized. In short, online education can no longer be regarded by researchers and practitioners as a uniform concept. This study, and others like it, provide strong evidence that strategies for online instruction may be as different from each other as they are from traditional classroom instruction in terms of effects on key student outcomes. Instead, online instructional strategies should be regarded as educational tools that may be combined with classroom instruction and with each other, as needed, to meet the needs of any educational circumstance.

Moreover, it is time to discard the outdated notion that online education is somehow wholly separate and distinct from traditional classroom instruction. The college educators who

will be most effective in the future will consider all available instructional strategies and will utilize empirical evidence to evaluate a number of possible combinations of strategies when designing a course. This study's results strongly suggest that there is no single best strategy of instruction (including classroom) for maximizing all possible student outcomes. Instead of being seen as disruptive innovations or as replacements for traditional classroom education, strategies for online instruction, just like different classroom teaching techniques, are simply tools for educators to use in creating optimal educational experiences for their students. If the research can keep pace with new developments in online instruction, then it will be possible for educators to consider all instructional strategies equally on their merits, without the need for arbitrary distinctions between what is "traditional" education versus "online" education. In spite of their differences, most valid instructional strategies have certain potential advantages over others, depending on the context. It behooves today's educators to consider the contexts of their courses and to make use of every tool at their disposal to educate their students.

Appendices

Appendix A
Demographic Information Survey

Age:
- ❍ 17 or younger
- ❍ 18-19
- ❍ 20-21
- ❍ 22-25
- ❍ 26 or older

Gender:
- ❍ Male
- ❍ Female

Racial background:
- ❍ White/Caucasian
- ❍ African American/Black
- ❍ American Indian/Alaskan Native
- ❍ Asian American/Asian
- ❍ Native Hawaiian/Pacific Islander
- ❍ Latino/Hispanic

Is English your native language?
- ❍ Yes
- ❍ No

Undergraduate major or field of study:
- ❍ Science, Technology, Engineering, or Math
- ❍ Other (Humanities, Social Science, Arts, etc.)
- ❍ Undecided

What is your GPA at <institution>?
- ❍ Less than 2.50
- ❍ 2.50-2.99
- ❍ 3.00-3.49
- ❍ 3.50-4.00+

Have you ever taken an online class?

- ❍ Yes
- ❍ No

Have you ever taken a class that taught you about intellectual property law?

- ❍ Yes
- ❍ No

Regardless of your answer to the previous question, how interested are you in the subject of intellectual property law?

- ❍ not at all interested
- ❍ slightly interested
- ❍ moderately interested
- ❍ very interested
- ❍ extremely interested

Appendix B
Post-Lesson Assessment of Comprehension

1. Only one of the following types of intellectual property is NOT specifically protected & mentioned in the U.S. Constitution? Which one?
    a. Patents
    b. Copyrights
    c. Trademarks

2. Architectural plans are most likely to fall under which of the following classifications of intellectual property?
    a. Patent
    b. Copyright
    c. Trademark

3. This is the only supranational body where you can file a consolidated patent application for all member countries – name the organization.
    Text Entry Answer: *The European Union*

4. The fees you pay to "rent" the rights to my intellectual property – what are they called? (please answer as a single word)
    Text Entry Answer: *Royalties*

5. Australia can use the word "champagne" on its sparkling wine but New Zealand cannot. That is because the word champagne is used on a widespread basis in Australia but not in New Zealand. What is the doctrine at play here?
    a. Grey Market Goods
    b. Generic Use
    c. Field of Use Restrictions

Questions 6 & 7 refer to the following information: Diego sells rum in the U.S. under the trademark "Havana Club" since 1990, but Maria actually registers the trademark in 2007, before Diego?

6. Who has superior rights – Diego or Maria?
    a. Diego
    b. Maria

7. Under what Doctrine does Diego have that right?
    Text Entry Answer: *First Use Doctrine*

8. Barbie Dolls have that unique face – is this a trademark or a copyright?
    a. *Trademark*
    b. *Copyright*

9. Which of the following is NOT a disadvantage or weakness of the Paris Convention?
    a. *No enforcement mechanisms*
    b. *Allows foreign and domestic IP holders to be treated differently*
    c. *No minimum standards required*

10. The Fair Use Doctrine allows unauthorized use of an IPR for certain purposes such as education, for example. Name another purpose for which the Fair Use Doctrine allows unauthorized use of an IPR.
    a. *Commentary*
    b. *News*
    c. *Humor or Parody*
    d. *All of the above*

11. If you plan to license the intellectual property of a foreign company in order to use it in the U.S., what step must you also take?
    a. *Gain the pre-approval of the U.S. government*
    b. *Notify the U.S. government of the licensing by registering the license*
    c. *You do not need to take any additional steps that are not specified in your contract with the IP holder.*

Questions 12 & 13 refer to the following information: Your small internet company is represented by a unique symbol. You have never bothered to register the trademark for the symbol, but you have been using this symbol in all of your online business operations.

12. Suppose you find out that a Brazilian retail company has recently begun using your company's symbol on their products. Can you successfully pursue legal recourse against the Brazilian company under the First Use Doctrine?
    a. *Yes*
    b. *No*

13. Now suppose that you discover that, over three years ago, one of your company's early employees actually registered the trademark for your company's symbol using the Madrid Protocol's centralized filing process. Does this new information provide you with a viable legal challenge to the Brazilian company's use of your symbol?
    a. *Yes*
    b. *No*

14. One primary reason that the TRIPS Agreement is so significant is because it makes the concepts of the Paris and Berne Conventions enforceable by what?
    a. *The Patent Cooperation Treaty*
    b. *The Geneva Act*
    c. *The World Trade Organization*
    d. *The Doha Declaration*

15. Which of the following patents is most likely to be held up as valid according to the U.S. Federal Patent Act?
    a. *A piñata made of standard glow-in-the-dark material*
    b. *A unique chemical compound that makes people sneeze*
    c. *A car with five wheels*
    d. *A newly-discovered tropical plant with significant cancer-fighting properties.*

Appendix C
Post-Lesson Questionnaire

Subject Matter

1. How interesting did you find the subject matter of the lesson (intellectual property law)?
    1. not at all interesting
    2. slightly interesting
    3. moderately interesting
    4. very interesting
    5. extremely interesting
2. How challenging did you find the subject matter of the lesson?
    1. not at all challenging
    2. slightly challenging
    3. moderately challenging
    4. very challenging
    5. extremely challenging
3. How useful do you think the knowledge you gained from the lesson will be to you in the future?
    1. not at all useful
    2. slightly useful
    3. moderately useful
    4. very useful
    5. extremely useful

Instructor

4. To what extent was the professor knowledgeable about the subject matter of the lesson?
    1. not at all knowledgeable
    2. slightly knowledgeable
    3. moderately knowledgeable
    4. very knowledgeable
    5. extremely knowledgeable
5. To what extent was the professor's presentation clear and well-organized?
    1. not at all clear and well-organized
    2. slightly clear and well-organized
    3. moderately clear and well-organized
    4. very clear and well-organized
    5. extremely clear and well-organized
6. To what extent did the professor make the material fun and engaging?
    1. not at all fun and engaging
    2. slightly fun and engaging
    3. moderately fun and engaging
    4. very fun and engaging
    5. extremely fun and engaging

Lesson Format

7. To what extent was the format (*student's specific treatment*) of the lesson appropriate for developing a solid understanding of the material?
    1. not at all appropriate
    2. slightly appropriate
    3. moderately appropriate
    4. very appropriate
    5. extremely appropriate
8. To what extent did the format of the lesson create a captivating learning experience?
    1. not at all captivating
    2. slightly captivating
    3. moderately captivating
    4. very captivating
    5. extremely captivating
9. To what extent did the format of the lesson make it straightforward and easy for you to fulfill your obligations as a student?
    1. not at all straightforward and easy
    2. slightly straightforward and easy
    3. moderately straightforward and easy
    4. very straightforward and easy
    5. extremely straightforward and easy

Interactions

10. How satisfied were you with your ability to interact with the instructor?
    1. not at all satisfied
    2. slightly satisfied
    3. moderately satisfied
    4. very satisfied
    5. extremely satisfied
11. How satisfied were you with your ability to interact with the other students in the class?
    1. not at all satisfied
    2. slightly satisfied
    3. moderately satisfied
    4. very satisfied
    5. extremely satisfied
12. Were all of your questions about the content of the lesson answered in the end?
    1. none were answered
    2. most were unanswered
    3. some were answered
    4. most were answered
    5. all were answered

Overall Learning Experience

13. Do you think you learned more or less in this lesson compared to a typical hour-long class session from your undergraduate studies?
    1. much less
    2. somewhat less
    3. about the same
    4. somewhat more
    5. much more
14. How likely would you be to take a full elective course on Intellectual Property Law if it were to be offered by this instructor in a similar format of instruction?
    1. extremely unlikely
    2. unlikely
    3. about a 50-50 chance
    4. likely
    5. extremely likely
15. How satisfied are you overall with this learning experience?
    1. not at all satisfied
    2. slightly satisfied
    3. moderately satisfied
    4. very satisfied
    5. extremely satisfied

Appendix D
Discussion Evaluation Rubric

Unacceptable (1)	Developing (2)	Proficient (3)	Exemplary (4)
Comment is uninformative, lacking in appropriate terminology, and bears little relation to the topic. Heavy reliance on opinion & personal taste, e.g., “I love it”, “I hate it”, “It’s bad” etc.	Comment is somewhat constructive, with signs of insight, but does not advance the conversation. Comment may be repetitive or not wholly relevant. Student does not use appropriate terminology.	Comment is fairly insightful & constructive; mostly uses appropriate terminology. Some support is provided, at least in general ways, to make connections between the topic and the student's comment. Clarification questions may be asked.	Comment is fully insightful, constructive, and uses appropriate terminology. Comment enhances lesson or discussion: they may ask a key question, elaborate, bring in relevant personal knowledge, move the discussion along, identify issues, or take the discussion to another level.

Adapted from Chang (Rubric for Assessing Student Participation) and Regina Public Schools (2003).

Appendix E
Quiz Score Regression

Regression Statistics	
Observations	303
Multiple R	0.430
R Square	0.185
Adjusted R Square	0.139

	Coefficients	Standardized Coefficients	p-values
Intercept	81.473		0.000
Group S	-15.887	-0.354	0.000
Group H	-7.036	-0.141	0.037
Group V	-12.711	-0.290	0.000
Group T	-6.897	-0.162	0.022
18-19 years old	2.675	0.072	0.230
>19 years old	-0.308	-0.007	0.915
Female	-5.357	-0.133	0.016
Asian	-1.024	-0.028	0.679
URM	-1.682	-0.035	0.606
English 2nd Language	-5.903	-0.154	0.007
STEM Major	0.460	0.013	0.820
3.00-3.49 GPA	-2.753	-0.072	0.356
>3.50 GPA	-0.937	-0.026	0.745
Online Experience	-2.324	-0.063	0.259
IP Law Experience	4.080	0.065	0.252
Moderate+ Interest	3.844	0.100	0.068

Appendix F
Participation Rates Logistic Regression

Regression Statistics				
Observations	303			
P-value	0.000			
		Coefficients	Standardized Coefficients	p-values
Intercept		-0.823		0.004
B		0.856	0.862	0.025
C		0.435	0.431	0.288
D		1.783	1.606	0.000
E		1.021	1.022	0.006

Appendix G
Comment Score Multiple Regression

Regression Statistics				
Observations	376			
Multiple R	0.541			
R Square	0.292			
Adjusted R Square	0.285			
		Coefficients	Standardized Coefficients	p-values
Intercept		2.100		0.000
B		-0.835	-0.508	0.000
C		0.376	0.245	0.125
D		0.516	0.468	0.006
E		0.412	0.378	0.029

Appendix H
Full Survey Regression

Regression Statistics	
Observations	303
Multiple R	0.482
R Square	0.232
Adjusted R Square	0.189

	Coefficients	Standardized Coefficients	p-values
Intercept	3.575		0.000
Group S	-0.166	-0.119	0.081
Group H	0.011	0.007	0.914
Group V	-0.440	-0.323	0.000
Group T	-0.546	-0.414	0.000
18-19 years old	0.132	0.114	0.050
>19 years old	-0.169	-0.116	0.053
Female	-0.012	-0.010	0.857
Asian	0.028	0.025	0.703
URM	0.119	0.079	0.227
English 2nd Language	0.185	0.155	0.005
STEM Major	0.059	0.053	0.331
3.00-3.49 GPA	-0.039	-0.033	0.663
>3.50 GPA	-0.106	-0.095	0.219
Online Experience	0.038	0.033	0.543
IP Law Experience	-0.065	-0.033	0.543
Moderate+ Interest	0.040	0.034	0.523

Appendix I
Survey Subject Questions Regressions

Regression Statistics	
Observations	303
Multiple R	0.438
R Square	0.192
Adjusted R Square	0.147

	Coefficients	Standardized Coefficients	p-values
Intercept	2.674		0.000
Group S	0.033	0.023	0.744
Group H	0.172	0.108	0.108
Group V	-0.203	-0.144	0.038
Group T	-0.187	-0.137	0.050
18-19 years old	0.043	0.036	0.540
>19 years old	-0.145	-0.097	0.116
Female	0.102	0.079	0.148
Asian	0.282	0.238	0.000
URM	0.200	0.129	0.055
English 2nd Language	0.275	0.223	0.000
STEM Major	0.003	0.003	0.964
3.00-3.49 GPA	-0.112	-0.091	0.240
>3.50 GPA	-0.193	-0.167	0.036
Online Experience	0.128	0.108	0.052
IP Law Experience	-0.068	-0.034	0.547
Moderate+ Interest	0.109	0.089	0.104

Appendix J
Survey Instructor Questions Regression

Regression Statistics	
Observations	303
Multiple R	0.517
R Square	0.268
Adjusted R Square	0.227

	Coefficients	Standardized Coefficients	p-values
Intercept	4.067		0.000
Group S	-0.227	-0.158	0.019
Group H	-0.022	-0.013	0.834
Group V	-0.416	-0.295	0.000
Group T	-0.722	-0.528	0.000
18-19 years old	0.122	0.102	0.074
>19 years old	-0.009	-0.006	0.922
Female	0.044	0.034	0.519
Asian	-0.103	-0.087	0.174
URM	0.142	0.092	0.156
English 2nd Language	0.102	0.082	0.130
STEM Major	0.107	0.092	0.084
3.00-3.49 GPA	0.009	0.008	0.919
>3.50 GPA	0.044	0.038	0.619
Online Experience	-0.039	-0.033	0.536
IP Law Experience	0.023	0.011	0.836
Moderate+ Interest	0.032	0.026	0.616

Appendix K
Survey Format Questions Regression

Regression Statistics	
Observations	303
Multiple R	0.360
R Square	0.130
Adjusted R Square	0.081

	Coefficients	Standardized Coefficients	p-values
Intercept	3.758		0.000
Group S	-0.290	-0.151	0.037
Group H	-0.196	-0.093	0.183
Group V	-0.510	-0.272	0.000
Group T	-0.549	-0.303	0.000
18-19 years old	0.198	0.125	0.043
>19 years old	-0.197	-0.099	0.121
Female	-0.035	-0.020	0.721
Asian	-0.004	-0.003	0.968
URM	0.062	0.030	0.666
English 2nd Language	0.145	0.089	0.131
STEM Major	0.138	0.090	0.120
3.00-3.49 GPA	0.026	0.016	0.841
>3.50 GPA	-0.101	-0.066	0.422
Online Experience	-0.064	-0.041	0.475
IP Law Experience	-0.058	-0.021	0.711
Moderate+ Interest	0.008	0.005	0.927

Appendix L
Survey Interactions Questions Regression

Regression Statistics	
Observations	303
Multiple R	0.446
R Square	0.199
Adjusted R Square	0.154

	Coefficients	Standardized Coefficients	p-values
Intercept	3.810		0.000
Group S	-0.044	-0.021	0.759
Group H	0.248	0.110	0.104
Group V	-0.625	-0.313	0.000
Group T	-0.539	-0.279	0.000
18-19 years old	0.088	0.052	0.383
>19 years old	-0.192	-0.091	0.143
Female	0.068	0.037	0.495
Asian	-0.166	-0.099	0.139
URM	0.030	0.014	0.840
English 2nd Language	0.056	0.032	0.571
STEM Major	0.066	0.040	0.471
3.00-3.49 GPA	-0.198	-0.115	0.142
>3.50 GPA	-0.207	-0.127	0.112
Online Experience	0.129	0.078	0.165
IP Law Experience	-0.107	-0.037	0.508
Moderate+ Interest	-0.062	-0.035	0.517

Appendix M
Survey Learning Experience Questions Regression

Regression Statistics	
Observations	303
Multiple R	0.435
R Square	0.189
Adjusted R Square	0.144

	Coefficients	Standardized Coefficients	p-values
Intercept	3.535		0.000
Group S	-0.281	-0.131	0.062
Group H	-0.148	-0.062	0.355
Group V	-0.435	-0.208	0.003
Group T	-0.736	-0.363	0.000
18-19 years old	0.204	0.115	0.054
>19 years old	-0.285	-0.128	0.039
Female	-0.230	-0.120	0.030
Asian	0.139	0.079	0.239
URM	0.159	0.069	0.307
English 2nd Language	0.355	0.194	0.001
STEM Major	-0.004	-0.002	0.969
3.00-3.49 GPA	0.090	0.050	0.525
>3.50 GPA	-0.070	-0.041	0.608
Online Experience	0.041	0.024	0.675
IP Law Experience	-0.126	-0.042	0.456
Moderate+ Interest	0.116	0.063	0.247

Appendix N
Completion Rates Logistic Regression

Regression Statistics	
Observations	425
P-value	0.021

	Coefficients	Standardized Coefficients	p-values
Intercept	0.896		0.090
Group S	0.046	0.026	0.896
Group H	-0.761	-0.428	0.022
Group V	0.329	0.185	0.352
Group T	0.768	0.431	0.044
18-19	0.054	0.037	0.833
>19	0.045	0.025	0.890
Female	-0.179	-0.111	0.495
Asian	0.467	0.323	0.089
URM	0.361	0.188	0.320
2nd Lang	-0.329	-0.218	0.186
STEM	0.064	0.045	0.780
3.00-3.49	-0.032	-0.021	0.928
>3.50	-0.030	-0.021	0.929
Online Exp.	-0.550	-0.384	0.020
IP Law Exp.	-0.183	-0.074	0.639
Moderate+	0.215	0.143	0.371

References

Agresti, A., & Finlay, B. (2009). *Statistical Methods for the Social Sciences, 4th ed.* Upper Saddle River, NJ: Prentice Hall.

Allen, I. E., & Seaman, J. (2009). *Learning on Demand: Online Education in the United States.* Newburyport, MA: Sloan Consortium.

Allen, I. E., & Seaman, J. (2011). *Going the Distance: Online Education in the United States, 2011.* Newburyport, MA: Sloan Consortium.

Allen, I. E., & Seaman, J. (2013). *Changing Course: Ten Years of Tracking Online Education in the United States.* Newburyport, MA: Sloan Consortium.

Allen, I. E., & Seaman, J. (2014). *Grade Change: Tracking Online Education in the United States.* Newburyport, MA: Sloan Consortium.

Allen, M., Bourhis, J., Burrell, N., & Mabry, E. (2002). Compariing Student Satisfaction with Distance Education to Traditional Classrooms in Higher Education: A Meta-Analysis. *American Journal of Distance Education*, 16(2), 83-97.

Al-Shalchi, O. N. (2009). The Effectiveness and Development of Online Discussions. *MERLOT Journal of Online Learning and Teaching*, 5(1), 104-108.

Amlie, T. (2003). Self-Selection Bias In On-Line Accounting Courses: Student Characteristics. *Journal of Business & Economics Research*, 1(8), 107-111.

Anderson, T. (2008). *The Theory and Practice of Online Learning.* Athabasca University Press.

Angus, S., & Watson, J. (2009). Does regular online testing enhance student learning in the numerical sciences? Robust evidence from a large data set. *British Journal of Educational Technology*, 40(2), 255-272.

Arizona Learning Systems. (1998). *Preliminary cost methodology for distance learning.* The State Board of Directors for Community Colleges of Arizona.

Aslanian, C., & Clinefelter, D. (2013). *Online College Students 2013: Comprehensive Data on Demands and Preferences.* Louisville, KY: The Learning House, Inc.

Ausburn, L. J. (2004). Course design elements most valued by adult learners in blended online education environments: an American perspective. *Educational Media International*, 41 (4), 327-337.

Bean, J. C., & Peterson, D. (1998). Grading Classroom Participation. *New Directions for Teaching and Learning*, 74, 33-40.

Benoit, P., Benoit, W., Milyo, J., & Hansen, G. (2006). *The effects of traditional vs. web-assisted instruction on student learning and satisfaction.* Report published by the University of Missouri.

Bernard, R., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., . . . Huang, B. (2004). How Does Distance Education Compare with Classroom Instruction? A Meta-Analysis of the Empirical Literature. *Review of Educational Research*, 74(3), 379–439.

Berrett, D. (2012, February 19). How 'Flipping' the Classroom Can Improve the Traditional Lecture. *The Chronicle of Higher Education*.

Billington, P. J., & Fronmueller, M. P. (2013). MOOCs and the Future of Higher Education. *Journal of Higher Education Theory & Practice*, 36-43.

Bishop, J. L., & Verleger, M. A. (2013). The Flipped Classroom: A Survey of the Research. *2013 ASEE Annual Conference.*

Boettcher, J. (2004). Online Course Development: What Does It Cost? *Campus Technology*, 17(12) 26-34.

Bowen, W. G., Chingos, M. M., Lack, K. A., & Nygren, T. I. (2012). *Interactive Learning Online at Public Universities: Evidence from Randomized Trials.* Ithaka S+R.

Brady, L. (2001). Fault Lines in the Terrain of Distance Education. *Computers and Composition*, 18(4), 347–358.

Bruner, J. S. (1966). *Toward a Theory of Instruction.* Harvard University Press.

Bullen, M. (1998). Participation and Critical Thinking in Online University Distance Education. *The Journal of Distance Education*, 13 (2), 1-32.

Burnett, C. (2003). Learning to Chat: Tutor participation in synchronous online chat. *Teaching in Higher Education*, 8(2), 247-261.

Byoun, S., & Zoljargal, D. (2012). Can Technology Free Higher Education From the Traditional Classroom. *Paper presented at the Association for Higher Education Research annual conference*, Las Vegas, NV.

Byrd, J., & Mixon, P. (2012). Revenue and E-Learning: do universities need an online presence. *Journal of Higher Education Policy and Management*, 34(6), 601-609.

Byrne, J. A. (2014, July 15). *Can Half The Business Schools Really Go Out Of Business?* Retrieved from www.poetsandquants.com: http://poetsandquants.com/2014/07/15/can-half-the-business-schools-really-go-out-of-business/

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and Quasi-Experimental Designs for Research.* Chicago: Rand McNally & Company.

Campbell, J. P., DeBlois, P. B., & Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. *Educause Review*, 42(4), 40-57.

Carey, K. (2009, April 3). What Colleges Should Learn From Newspapers' Decline. *The Chronicle of Higher Education*.

Carnoy, M., Rabling, B. J., Castano-Munoz, J., Montoliu, J. M., & Sancho-Vinuesa, T. (2012). Who attends and completes virtual universities: the case of the open University of Catalonia (UOC). *Higher Education*, 63(1), 53-82.

Caspi, A., Chajut, E., & Saporta, K. (2008). Participation in class and in online discussions: Gender differences. *Computers & Education*, 50(3), 718-724.

Centra, J. A. (1993). *Reflective Faculty Evaluation: Enhancing Teaching and Determining Faculty Effectiveness.* San Francisco: Jossey-Bass.

Chang, T. (n.d.). *Rubric for Assessing Student Participation.* Retrieved from Eberly Center for Teaching Excellence at Carnegie Mellon University: http://www.cmu.edu/teaching/assessment/examples/courselevel-bytype/performancecriteria/course_rubricparticipation-ArtSociety.html

Choi, H., & Johnson, S. (2005). The Effect of Context-Based Video Instruction on Learning and Motivation in Online Courses. *American Journal of Distance Education*, 19(4), 215-227.

Christensen, C. (1997). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail.* Harvard Business Review Press.

Christensen, C. M., & Eyring, H. J. (2011). *The Innovative University: Changing the DNA of higher education from the inside out.* John Wiley & Sons.

Clark, R. E. (1983). Reconsidering Research on Learning from Media. *Review of Educational Research*, 53(4), 445-459.

Cleveland-Innes, M., & Ally, M. (2013). Affective Learning Outcomes in Workplace Training: A Test of Synchronous vs. Asynchronous Online Learning Environments. *Canadian Journal of University Continuing Education*, 30(1).

Coates, H., James, R., & Baldwin, G. (2005). A Critical Examination Of The Effects Of Learning Management Systems On University Teaching And Learning. *Tertiary Education and Management*, 11(1), 19-36.

Cole, M. T., Shelley, D. J., & Swartz, L. B. (2012). Academic Integrity and Student Satisfaction in an Online Environment. In H. Yang, & W. S., *Cases on Online Learning Communities and*

Beyond: Investigations and Applications (pp. 1-19). Hershey, PA: Information Science Reference.

Conrad, R., & Donaldson, J. A. (2011). *Engaging the Online Learner: Activities and Resources for Creative Instructio.* John Wiley & Sons.

Couper, M. P., & Bosnjak, M. (2010). Internet Surveys. In P. V. Marsden, & J. D. Wright, *Handbook of Survey Research, 2nd Ed.* (pp. 527-550). Emerald Group Publishing.

Coursera Blog. (2012, August 9). Retrieved from Coursera website: http://blog.coursera.org/post/29062736760/coursera-hits-1-million-students-across-196-countries

Creswell, J. W. (2009). *Research Design, 3rd Edition.* Thousand Oaks, CA: SAGE Publications.

Cuban, L. (1993). *How Teachers Taught: Constancy and Change in American Classrooms, 1890-1990.* New York: Teachers College Press.

Cuban, M. (2012, May 13). *The Coming Meltdown in College Education & Why The Economy Won't Get Better Any Time Soon*. Retrieved from www.blogmaverick.com: http://blogmaverick.com/2012/05/13/the-coming-meltdown-in-college-education-why-the-economy-wont-get-better-any-time-soon/

de Vaus, D. A. (1991). *Surveys in Social Research.* Allen & Unwin.

Dewey, J. (1938). *Experience and Education.* New York: Macmillan.

Diaz, D. P. (2002, May/June). Online Drop Rates Revisited. *The Technology Source*. Retrieved from http://technologysource.org/article/online_drop_rates_revisited/

Dommeyer, C. J., Baum, P., W., H. R., & Chapman, K. S. (2004). Gathering faculty teaching evaluations by in-class and online surveys: their effects on response rates and evaluations. *Assessment & Evaluation in Higher Education*, 29(5), 611-623.

Drago, W. A., & Wagner, R. J. (2004). Vark preferred learning styles and online education. *Management Research News*, 27(7), 1-13.

Dunn, R. (1983). Can Students Identify Their Own Learning Styles? *Educational Leadership*, 40(5), 60-62.

Eberle, J., & Childress, M. (2005). Using heutagogy to address the needs of online learners. *Encyclopedia of distance learning*, 4, 1945-1951.

Ehrenberg, R. (2012). American Higher Education in Transition. *The Journal of Economic Perspectives*, 26(1), 193-216.

Elvers, G. C., Polzella, D. J., & Graetz, K. (2003). Procrastination in online courses: Performance and attitudinal differences. *Teaching of Psychology*, 30(2), 159-162.

Ely, D. P. (2003). Selecting Media for Distance Education. *ERIC Digest*.

Eom, S., Wen, J., & Ashill, J. (2006). The Determinants of Students' Perceived Learning Outcomes and Satisfaction in University Online Education: An Empirical Investigation. *Decision Sciences Journal of Innovative Education*, 4(2), 215-235.

Ernst, J. V. (2008). A Comparison of Traditional and Hybrid Online Instructional Presentation in Communication Technology. *Journal of Technology Education*, 19(2).

Figlio, D. N., Rush, M., & Yin, L. (2010). Is it Live or is it Internet? Experimental Estimates of the Effects of Online Instruction on Student Learning (No. w16089). *National Bureau of Economic Research*.

Fitzpatrick, M. (2012, June 24). Classroom Lectures Go Digital. *The New York Times*.

Fowler, F. J. (2002). *Survey Research Methods, 3rd Ed.* Sage Publications.

Fowler, F. J. (2013). *Survey Research Methods, Fifth Edition.* Sage.

Fox, J. (2010, March 3). *What does a corporate web video cost? 25 Factors (with prices) that affect corporate video production costs.* Retrieved from www.onemarketmedia.com: http://www.onemarketmedia.com/2010/03/03/what-does-a-web-video-cost-25-factors-with-prices-that-affect-video-production-costs/

Fritz, J. (2011). Classroom walls that talk: Using online course activity data of successful students to raise self-awareness of underperforming peers. *The Internet and Higher Education*, 14(2), 89-97.

Gaytan, J., & McEwen, B. C. (2007). Effective Online Instructional and Assessment Strategies. *American Journal of Distance Education*, 21(3), 117-132.

Ge, Z. (2012). Cyber Asynchronous versus Blended Cyber Approach in Distance English Learning. *Journal of Educational Technology & Society*, 15(2), 286-297.

Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education*, 57(4), 2333-2351.

Glass, G. V., & Stanley, J. C. (1970). *Statistical Methods in Education and Psychology.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Gleason, B. J. (2004). Retention Issues in Online Programs: A Review of the Literature. *A paper presented at the Second AIMS International Conference on Management*, (pp. 395-398). Calcutta, India.

Gorsky, P., & Caspi, A. (2005). A Critical Analysis of Transactional Distance Theory. *Quarterly Review of Distance Education*, 6(1), 1-11.

Gunawardena, C., & McIsaac, M. S. (2003). Distance Education. In D. H. Jonassen, & M. Driscoll, *Handbook of Research on Educational Communications and Technology* (pp. 355-396). Routledge.

Guo, P. J., Kim, J., & Rubin, R. (2014). How Video Production Affects Student Engagement: An Empirical Study of MOOC Videos. *Proceedings of the first ACM conference on Learning@ scale conference.* ACM.

Hardesty, L. (2013, March 6). Retrieved April 10, 2014, from MIT News: http://newsoffice.mit.edu/2013/edx-summit-0306

Hardy, N. (2003). Online Ratings: Fact and Fiction. In D. L. Sorenson, & T. D. Johnson, *Online Student Ratings of Instruction: New Directions for Teaching and Learning* (pp. 31-38). John Wiley & Sons.

Harman, K., & Koohang, A. (2005). Discussion Board: A Learning Object. *Interdisciplinary Journal of E-Learning and Learning Objects*, 1(1), 67-77.

Hase, S., & Kenyon, C. (2000). From Andragogy to Heutagogy. *Ultibase Articles*, 5(3), 1-10.

Helms, S. A. (2012). Blended/hybrid courses: a review of the literature and recommendations. *Interactive Learning Environments*, 1-7.

Ho, A. D., Reich, J., Nesterko, S., Seaton, D. T., Mullaney, T., Waldo, J., & Chuang, I. (2014). HarvardX and MITx: The First Year of Open Online Courses, Fall 2012-Summer 2013. *(HarvardX and MITx Working Paper No. 1)*.

Hoffman, K. M. (2003). Online Course Evaluation and Reporting in Higher Education. In D. L. Sorenson, & T. D. Johnson, *Online Student Ratings of Instruction: New Directions for Teaching and Learning* (pp. 25-29). John Wiley & Sons.

Howe, J. (2013, February 12). Clayton Christensen Wants to Transform Capitalism. *Wired Magazine*.

Howell, S. L., Williams, P. B., & Lindsay, N. K. (2003). Thirty-two Trends Affecting Distance Education: An Informed Foundation for Strategic Planning. *Online Journal of Distance Learning Administration*, 6(3).

Hrastinski, S. (2008a). The potential of synchronous communication to enhance participation in online discussions: A case study of two e-learning courses. *Information & Management*, 45(7), 499-506.

Hrastinski, S. (2008b). Asynchronous and Synchronous E-Learning. *EDUCAUSE Quarterly*, 31(4).

Jahng, N., Krug, D., & Zhang, Z. (2007). Student achievement in online distance education compared to face-to-face education. *European Journal of Open, Distance and E-Learning*, Retrieved 3.18.2013 from http://www.eurodl.org/materials/contrib/2007/Jahng_Krug_Zhang.htm.

Johnson, G. M. (2006). *Synchronous and Asynchronous Text-Based CMC in Educational Contexts: A Review of Recent Research.* 50(4), 46-53: TechTrends.

Johnson, L., S., A. B., Estrada, V., & Freeman, A. (2014). *NMC Horizon Report: 2014 Higher Education Edition.* Austin, TX: The New Media Consortium.

Johnson, S., Aragon, S., & Shaik, N. (2000). Comparative Analysis of Learner Satisfaction and Learning Outcomes in Online and Face-to-Face Learning Environments. *Journal of Interactive Learning Research*, 11(1), 29-49.

Joy, E. H., & Garcia, F. E. (2000). Measuring Learning Effectiveness: A New Look at No-Significant-Difference Findings. *The Journal of Asynchronous Learning Networks*, 4(1), 33-39.

Jung, I. (2001). Building a Theoretical Framework of Web-Based Instruction in the Context of Distance Education. *British Journal of Education Technology*, 1-17.

Kearsley. (1997). *A Guide to Online Education*. Retrieved from http://w3.cs.com.uy/u/gux/doc/tech/online.htm on 4/10/2014

Keengwe, J., & Kidd, T. T. (2010). Towards Best Practices in Online Learning and Teaching in Higher Education. *MERLOT Journal of Online Learning and Teaching*, 6(2), 533-541.

Keller, B. (2011, October 2). The University of Wherever. *The New York Times*.

Keppel, G. (1991). *Design and Analysis.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Khare, A., & Lam, H. (2008). Assessing Student Achievement and Progress with Online Examinations: Some Pedagogical and Technical Issues. *International Journal on E-Learning*, 7(3), 383-402.

Kirk, R. (1982). *Experimental Design.* John Wiley & Sons.

Kirkwood, A., & Price, L. (2008). Assessment and student learning: a fundamental relationship and the role of information and communication technologies. *Open Learning*, 23(1), 5-16.

Knowles, M. (1973). *The Adult Learner: A Neglected Species.* Houston, TX: Gulf Publishing Company.

Kozma, R. (1994). Will media influence learning? Reframing the debate. *Educational Technology Research and Development*, 42(2), 7-19.

Larreamendy-Joerns, J., & Leinhardt, G. (2006). Going the Distance With Online Education. *Review of Educational Research*, 76(4), 567–605.

Laurillard, D. (2013). *Rethinking University Teaching: A Conversational Framework for the Effective Use of Learning Technologies.* Routledge.

Laws, R. D., Howell, S. L., & Lindsay, N. K. (2003). Scalability in Distance Education: "Can We Have Our Cake and Eat it Too?". *Online Journal of Distance Learning Administration*, 6(4).

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., . . . Wolff, S. (2009). A brief history of the Internet. *ACM SIGCOMM Computer Communication Review*, 39(5), 22-31.

Li, L., Finley, J., Pitts, J., & Guo, R. (2010). Which is a Better Choice for Student-Faculty Interaction: Synchronous or Asynchronous Communication? *Journal of Technology Research*, 2.

Lieblein, E. (2000). Critical factors for successful delivery of online programs. *The Internet and Higher Education*, 3(3), 161-174.

Lievrouw, L. A. (2001). Instructional media and the 'no significant difference' phenomenon. *ICA Newsletter*, 29(5), 5-6.

Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education*, 54(2), 588-599.

Maguire, L. (2005). Literature review–faculty participation in online distance education: Barriers and motivators. *Online Journal of Distance Learning Administration*, 8(1).

Marcel, K. (2002). Can Law be Taught Effectively Online? *JURIST*.

Matthews, D. (1999). The Origins of Distance Education and Its Use in the United States. *T H E Journal (Technological Horizons In Education)*, 27(2).

Mazur, E. (2009). Farewell, Classroom? *Science*, 50-51.

McDonald, J. (2002). Is "As Good As Face-to-Face" as Good as it Gets? . *Journal of Asynchronous Learning Networks*, 6(2), 10-23.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). *Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies.* U.S. Department of Education: Office of Planning, Evaluation, and Policy Development.

Meeker, M., & Wu, L. (2013). *2013 Internet Trends.* Kleiner Perkins Caufield & Byers, Technical Report.

Miller, K. H. (2004). The Law Catches Up with Distance Learning. *Technological Horizons in Education*, 31(7), 31-34.

Moore, M. (1993). Theory of Transactional Distance. In D. Keegan, *Theoretical Principles of Distance Education* (pp. 22-38). New York: Routledge.

Morris, L., Finnegan, C., & Wu, S. (2005). Tracking student behavior, persistence, and achievement in online courses. *The Internet and Higher Education*, 8(3), 221–231.

Morrow, B. V., & Shi, S. (2006). E-Conferencing for Instruction: What Works? *Educause Quarterly*, 29(4), 42.

Moser, C. A. (1958). *Survey Methods in Social Investigation.* William Heinemann.

Mundfrom, D. J., Perrett, J. J., Schaffer, J., Piccone, A., & Roozeboom, M. (2006). Bonferroni adjustments in tests for regression coefficients. *Multiple Linear Regression Viewpoints*, 32(1), 1-6.

Ng, K.-C. (2000). Costs and Effectiveness of Online Courses in Distance Education. *Open Learning: The Journal of Open, Distance and e-Learning*, 15(3), 301-308.

Ngwenya, J., Annand, D., & Wang, E. (2004). Supporting Asynchronous Discussions Among Online Learners. In T. Anderson, & F. Elloumi, *Theory and Practice of Online Learning* (pp. 319-347). Athabasca University.

Offir, B., Lev, Y., & Bezalel, R. (2008). Surface and Deep Learning Processes in Distance Education: Synchronous versus Ansynchronous Systems. *Computers & Education*, 51, 1172-1183.

Oztok, M., Zingaro, D., Brett, C., & Hewitt, J. (2013). Exploring asynchronous and synchronous tool use in online courses. *Computers & Education*, 60(1), 87-94.

Parsad, B., & Lewis, L. (2008). *Distance Education at Degree-Granting Postsecondary Institutions: 2006–07.* National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubs2009/2009044.pdf

Parthasarathy, M., & Smith, M. (2009). Valuing the Institution: An Expanded List of Factors Influencing Faculty Adoption of Online Education. *Online Journal of Distance Learning Administration*, 12(2).

Phipps, R., & Merisotis, J. (1999). *What's the Difference? A Review of Contemporary Research on the Effectiveness of Distance Learning in Higher Education.* Washington, D.C.: Institute for Higher Education Policy.

Piaget, J. (1929). *The child's conception of the world (Vol. 213).* Rowman & Littlefield.

Piaget, J. (1970). Piaget's Theory. In P. H. Mussen, *Carmichael's Manual of Child Psychology 3rd Edition* (pp. 803-832). New York: Wiley.

Power, M. (2008). The Emergence of a Blended Online Learning Environment. *Journal of Online Learning and Teaching*, 4(4), 503-514.

Regina Public Schools. (2003, October 1). *Class Participation Rubric.* Retrieved from Rubrics for Teachers: http://www.rubrics4teachers.com/participation.php

Reiser, R. (1994). Clark's invitation to the dance: An instructional Designer's response. *Educational Technology Research and Development*, 42(2), 45-48.

Rhoads, R. A., Berdan, J., & Toven-Lindsey, B. (2013). The Open Courseware Movement in Higher Education: Unmasking Power and Raising Questions about the Movement's Democratic Potential. *Educational Theory*, 63(1), 87-109.

Roblyer, M. D., & Wiencke, W. R. (2003). Design and Use of a Rubric to Assess and Encourage Interactive Qualities in Distance Courses. *American Journal of Distance Education*, 17(2), 77-98.

Romero, C., Ventura, S., & Garcia, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, 51(1), 368-384.

Rovai. (2004). A Constructivist Approach to Online College Learning. *Internet and Higher Education*, 7, 79-93.

Rovai, F., & Jordan, H. (2004). Blended Learning and Sense of Community: A Comparative Analysis with Traditional. *International Review of Research in Open and Distance Learning*, 5(2).

Rumble, G. (2001). The Costs and Costing of Networked Learning. *Journal of Asynchronous Learning Networks*, 5(2), 75-96.

Russell, T. L. (1999). *The No Significant Difference Phenomenon.* Raleigh, NC: North Carolina State University.

Saad, L., Busteed, B., & Ogisi, M. (2013, October 15). *In U.S., Online Education Rated Best for Value and Options.* Retrieved from Gallup Website: http://www.gallup.com/poll/165425/online-education-rated-best-value-options.aspx

Sendag, S., & Odabasi, H. F. (2009). Effects of an online problem based learning course on content knowledge acquisition and critical thinking skills. *Computers & Education*, 53(1), 132-141.

Shelley, D. J., Swartz, L. B., & Cole, M. T. (2007). A Comparative Analysis of Online and Traditional Undergraduate Business Law Classes. *International Journal of Information and Communication Technology Education*, 3(1), 10-21.

Sheridan, D., White, D., & Gardner, L. A. (2002). Cecil: the first web-based LMS. *ASCILITE*, 603-611.

Siemens, G. &. (2011). Penetrating the fog: Analytics in learning and education. *Educause Review*, 46(5), 30-42.

Simonson, M. (2003). Definition of the field. *Quarterly Review of Distance Education*, 4(1), 7-8.

Sitzmann, T., Kraiger, K., Stewart, D., & Wisher, R. (2006). The comparative effectiveness of web-based and classroom instruction: A meta-analysis. *Personnel Psychology*, 59(3), 623-664.

Skiba, D. (2012). Disruption in Higher Education: Massively Open Online Courses (MOOCs). *Nursing Education Perspectives*, 33(6), 416-417.

Skylar, A. A. (2009). A Comparison of Asynchronous Online Text-Based Lectures and Synchronous Interactive Web Conferencing Lectures. *Issues in Teacher Education*, 18(2), 69-84.

Smith, P. L., & Ragan, T. J. (1999). *Instructional Design.* New York: Wiley.

Swan, K. (2002). Building Learning Communities in Online Courses: the importance of interaction. *Education, Communication & Information*, 2(1), 23-49.

Taylor, J. C. (2001). Fifth Generation Distance Education. Retrieved from http://www.ascilite.org.au/ajet/e-jist/docs/vol4no1/Taylor.pdf.

Tshibalo, A. (2007). The potential impact of computer-aided assessment technology in higher education. *South African Journal of Higher Education*, 21, 684-693.

Tu, C., & Corry, M. (2003). Designs, Management Tactics, and Strategies in Asynchronous Learning Discussions. *The Quarterly Review of Distance Education*, 4(3), 303-315.

Tyler, R. W. (1949). *Basic Principles of Curriculum and Instruction.* Chicago: University of Chicago Press.

U.S. Department of Education, National Center for Education Statistics. (2012). Digest of Education Statistics. *(NCES 2012-001)*, Chapter 3.

Vilaseca, J., & Castillo, D. (2008). Economic efficiency of e-learning in higher education: An industrial approach. *Intangible Capital*, 4(3), 191-211.

Volery, T., & Lord, D. (2000). Critical Success Factors in Online Education. *International Journal of Educational Management*, 14(5), 216-223.

Vonderwell, S. (2003). An examination of asynchronous communication experiences and perspectives of students in an online course: a case study. *The Internet and Higher Education*, 6(1), 77-90.

Vonderwell, S., Liang, X., & Alderman, K. (2003). Asynchronous Discussions and Assessment in Online Learning. *Journal of Research on Technology in Education*, 39(3), 309-328.

Vrasidas, C. (2000). Constructivism versus objectivism: Implications for interaction, course design, and evaluation in distance education. *International Journal of Educational Telecommunications*, 6(4), 339-362.

Vygotsky, L. S. (1980). *Mind in Society: The Development of Higher Psychological Processes.* Harvard University Press.

Wallen, N. E., & Fraenkel, J. R. (2000). *Educational Research: A Guide To the Process.* Routledge.

Walsh, E., O'Keeffe, N., Delaney, L., Fox, S., Brunton, J., Costello, E., & Morrissey, A. (2012). Enhancing the teaching and learning experience of distance education through the use of synchronous online tutorials. Cyprus: EADTU Conference.

Wang, S., & Hsu, H. (2008). Use of the Webinar Tool (Elluminate) to Support Training: The Effects of Webinar-Learning Implementation from Student-Trainers' Perspective. *Journal of Interactive Online Learning*, 7(3), 175-194.

Ward, B. (2004). The Best of Both Worlds: A Hybrid Statistics Course. *Journal of Statistics Education*, 12(3), 74-79.

Xu, D., & Jaggars, S. (2011). The Effectiveness of Distance Education Across Virginia's Community Colleges: Evidence From Introductory College-Level Math and English Courses. *Educational Evaluation and Policy Analysis*, 33(3), 360–377.

Xu, D., & Jaggars, S. (2013a). Examining the Effectiveness of Online Learning Within a Community College System: An Instrumental Variable Approach. *CCRC Working Paper No. 56*.

Xu, D., & Jaggars, S. (2013b). Adaptability to Online Learning: Differences Across Types of Students and Academic Subject Areas. *CCRC Working Paper No. 54*.

Zhao, Y., Lei, J., Yan, B., Lai, C., & Tan, H. (2005). What makes the difference? A practical analysis of research on the effectiveness of distance education. *Teachers College Record*, 107 (8), 1836–1884.