COMPARING STUDENTS' LEARNING OUTCOMES AND SATISFACTION
IN ONLINE, HYBRID AND FACE-TO-FACE EDUCATION COURSES

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF EDUCATION

by
Lori M. Bailey
Diploma Date June 2020

Examining Committee Members:

Dr. Joseph DuCette, Advisory Chair, Psychological Studies in Education
Dr. Joan Shapiro, Policy, Organizational, and Leadership Studies
Dr. Steven Jay Gross, Policy, Organizational and Leadership Studies
Dr. Judith Stull, Policy, Organizational and Leadership Studies

ABSTRACT

The this study examined courses taught within a single College of Education over six semeters to compare face-to-face, hybrid, and online instructional methods as related to student achievement, persistence, and satisfaction. Additionally, a comparison of key student characteristics including sex, race/ethnicity, and residency status was conducted. This study extends the existing literature supporting "no meaningful significant difference" between instructional delivery methods by specifically focusing on courses of similar curriculum as offered within the specific discipline of educational studies. The results reinforce that administrators and instructors should continue to expand access to courses through the flexibility of online and hybrid learning. However, as programs expand their course offerings, further investigation is warranted into the barriers to hybrid and online learning for certain groups of educational studies students within this institution, including women, Asians, and out-of-state residents.

DEDICATION

For all my colleagues and professors who wouldn't let me give up; and for my family

who supported and encouraged me throughout.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

The current national conversation includes an intense debate over the causes, solutions, and repercussions of the ever-rising costs of undergraduate and graduate education in the U.S. A recent survey[i] of scholarly articles published from only as far back as January 2012 revealed over 50 entries classified in the subject area of "college costs," and included discussions on a broad range of financial and enrollment implications from the general student and institution (Dervarics, 2013; Evans, 2013; Grummon, 2010) to the specific major and profession (Elliott & Soo, 2013; Felton, 2013; Liu, 2013; Stone-Brown, 2013).

In his overview of "Costs and Productivity in Higher Education," William G. Bowen (2013), a former president of Princeton University, succinctly lays out the underlying causes for the "seemingly inexorable tendency for institutional costs per student . . . to rise faster than costs in general over the long term" (p. 2). Bowen provides a substantive overview of his and William J. Baumol's theory of "cost disease" that impacts "labor intensive industries such as the performing arts and education" where, due to the nature of the industry itself, productivity cannot easily be increased through technological advances, innovation, and other capital investment (p. 3). And, Bowen offers some sobering statistics to emphasize the undeniable realities around the rising costs of education (p. 3-4). He also includes a summary list of the "usual suspects" around efficiency and affordability, which have also been much discussed, and which continue to "push up educational costs" including:

---

[i] A search of Temple University Library databases was conducted using Temple's "Summon" inclusive search tool. The search term was only: "SubjectTerms:"college costs"" and results were limited to a date of publication from January 1, 2012 to the date of the search: April 20, 2014 and to "articles from scholarly publications, including peer-review" and included both journal and newspaper sourced articles. A total of 53 items were identified.

- inefficiencies: the inability of [Higher Education] institutions to subtract services and scope, once added, and to eliminate or modify "antiquated organizational structures such as 'centers' (p. 11);

- over-focus on rankings and reputation: whereby the need to enhance reputation leads to unsustainable policies and investment in research and activities that provide short-term gains;

- increasing time-to-degree: resulting from a "combination of declines in resources at these less selective public institutions and the tendency for students to work more hours for pay" (p. 16); and

- decreasing state appropriations: resulting in more burden on the individual student and more student loan debt across the system.

Taken together, these structural, social, and political forces portend significant changes for the American Higher Education (HE) system as it is currently experienced by the little over 20 million students enrolled in degree-granting institutions as of the Fall of 2017 (Ginder, Kelly-Reid, & Mann, 2018).

The promise of technology – specifically online learning – has emerged as a controversial instructional method in HE institution's aspiration to expand and improve teaching and learning, while simultaneously reducing costs. In "From Baumol's Cost Disease to Moore's Law: Bending the Cost Curve in Higher Education," Kamenetz (2013, pp. 9-26) provides a succinct accounting of the various structures and pressures contributing to the current and unsustainable trend of ever-increasing costs in HE. Like many others, Kamenetz (2013) ends her summary of the current woeful financial state of HE with a nod toward the potential to be found within technological innovation – technology that will be needed to deliver educational opportunities "better, faster, cheaper" (p. 23). Indeed, faced with a precarious future based on the traditional models of teaching and learning, HE looks to new delivery methods and media – specifically online learning and Distance Education (DE) – as a "disruptive technology" that will both

change its structure and improve its outcomes. At the same time, critics of online learning question whether technology-enabled modes of learning can live up to its promise to meet the needs of non-traditional learners and deliver better student learning than traditional education (Bonvillian & Singer,2013; Demirci, 2014; Keramidas, 2012; Wildavksy, 2013). And going beyond its promise of providing equitable learning opportunities while expanding access to education for diverse populations of students – all at a presumed lower cost – advocates of the expansion of online learning in HE point to its possibilities: to deliver instruction that simultaneously supports alternative learning styles, while providing experiences that are not otherwise possible in the traditional classroom (Alexander & Sonalini, 2013; McNatt, 2019; Plohn, 2014; Warner, 2016).

In *Disrupting Class: How Disruptive Innovation Will Change the Way the World Learns,* Christensen (2011) references Gardner's eight intelligences (p. 26) and makes a strong argument that, in the current, traditional classroom, a student's success largely depends if "an educational approach is [already] well aligned with one's strong intelligences or aptitudes" (p. 27).   In advocating for a new kind of learning that supports differentiated learning and addresses the needs of students outside those who already succeed in traditional HE classrooms, Christensen then presents a detailed exploration of and justification for a new approach to education that emphasizes individualized instruction for students; caters to different learning styles and methods; and avoids repeating the existing educational patterns that enable only those students who are already prepared to succeed in traditional learning environments. He compares the current approach to teaching and learning to that of a computer operating system, like Windows or Macintosh, that requires the user to adapt to the requirements of the system. In contrast, he argues, education *should* be about providing a curriculum that adapts learning techniques and strategies to different intelligences.   And, he posits a technology-enabled approach to teaching that is designed more like an application (or

app) environment that adapts to the user by providing customized "chunks" or "modules" of learning that can be connected together, as the user needs, to meet specific goals and tasks (p. 32). Instead of the traditional focus on "standardization" (p. 35) that delivers the same learning environment to all students, regardless of their different learning needs, Christensen envisions a "student-centered" model "that is completely modular" (p. 38) and allows students to adapt the curriculum to their specific needs. Recent advances in the technology, Christensen argues, make possible this new vision of HE learning systems capable of catering to individual student needs and abilities, while reducing costs and tearing down barriers to access for quality and affordable degrees and credentials. Technology, for Christensen, is the means by which HE institutions will deliver this idealized adaptive learning environment. He anticipates this revolutionary model will arrive soon, whether learning institutions are ready or not.

To Christensen's point, the number of online enrollments continues to grow at a rate that far exceeds the rate of overall enrollments in higher education (Allen & Seaman, 2014). According to an updated report compiled by Seaman, Allen and Seaman (2018), at the same time that the U.S. higher education institutes are experiencing an overall decline in the total number of students enrolled (p. 7), the number of students taking at least one or more distance education courses, including both graduate and undergraduate students, increased in each of the reported periods to almost 30% by 2015 (p. 11). In a recent summary of 2017 data from the U.S. Department of Education (Ginder, Kelly-Reid, & Mann, 2018), of the over 20 million students enrolled in Title IV institutions in Fall 2017, over three million were enrolled exclusively in distance education courses and another six million were enrolled in at least some distance education courses [table 3 page 10]. But, is online learning merely a continuation of the traditional models of teaching and learning replicated in virtual environments or is this truly "disruptive change"? Does more online teaching also mean more learning or even equitable learning? Or does the expansion of

online learning, as Christensen posited, open the door toward better, more effective, individualized learning?

Examining institutional perceptions of online learning Dhilla (2017) noted that several studies have questioned the efficacy of online courses, citing lower retention rates and unclear learning outcomes as cause for concern (p.136). *Changing Course: Ten Years of Tracking Online Education in the United States* (Allen & Seaman, 2013) reports the result of an annual survey conducted by the Babson Survey Research Group and the Quahog Research Group since 2002. The report focuses on the state of online learning in U.S. higher education and provides insights from the point of view of chief academic officers at 2,800 colleges and universities. In the period from 2002 to 2012, the annual survey reported a greater than 19% increase in chief academic leaders who said that online learning is critical to their institutions long-term strategy, along with a reported high in the most-recent survey of around 32% of students taking at least one online course. At the same time, only about 30% of chief academic officers reported believing "their faculty accept the value and legitimacy of online education" (p. 27), while almost 77% of those same leaders now considered online education as "the same or superior" to traditional classroom instruction (p. 24).

Although online and hybrid learning has gained much attention for its potential to deliver effective student learning at lower cost (increased efficiency), Bowen, Chingos, Lack, and Nygren (2013) point out that "there is very little rigorous evidence on learning outcomes for students receiving instruction online" (p.59). Institutions must assess not-only the value-added in decreasing the cost of instruction through online learning, but also its impact on student success. One way to evaluate the effectiveness and efficiency of student learning is by evaluating student achievement. Yet, studies focusing on the student's outcomes in comparing the traditional and online section of the same course question whether it is enough to evaluate student achievement via grades alone (Warren

& Holloman, 2005). Another approach is to evaluate student perception and satisfaction. At the same time, even though students may be equally satisfied with participation in DE, it is possible they may not be learning as much in that environment.

Student Evaluations of Teaching (SETs), also known as course evaluations, are one of the most ubiquitous and widely adopted measure of both teaching quality and student satisfaction. Often deployed systemically across institutions, SETs assess the quality of instruction at institutions of higher education and serve as a means of comparing teacher performance across multiple dimensions (Becker, Bosshardt, & Wattas, 2012). Often controversial in their varied application and interpretation, SETs have been used as evidence of instructional quality and student satisfaction for various purposes: including: faculty promotions (Uttl & Smibert, 2017), students' college applications (Alter & Reback, 2014) as well as students' course selection (Wilhelm, 2004) These applications and their varied implications have consequently sparked a large scientific literature on SETs resulting in a body of evidence that "have both supported and questioned the validity of SETs" (p. 684, Wright & Jenkins-Guarnier, 2012). Situated within this larger body of research, this study does not address areas of concern around biasing variables within student evaluations based on characteristics of the instructor such as sex, age, attractiveness, and other much-discussed factors (Baldwin & Blattner, 2003; Boehmer & Wood, 2017; Wood, Kiggins, & Kickham, 2017). Rather, this study accepts as proven the established relationship between accepted indicators of student learning and SETs, as demonstrated through meta-data analyses like those conducted by Wright and Jenkins-Guarneir (2012), whereby students who achieved more in a course also rated the teaching as more effective and are, therefore, "valid measures of instructor skill and teaching effectiveness" (p. 693). Further research has shown a variation of teaching evaluations as a function of discipline (Felton, Mitchell, & Stinson 2004; Kogan, 2014) generating studies focused on specific courses or specific fields such

as business, physiology, psychology, social work, business (Anderson & Krichbaum, 2017; Haywood & Murty, 2018; Secret, Bentley, & Kadolph, 2016; Weldy, 2018). Given the accepted value of using SETs as a quality indicator of student perceptions of learning, and, along with other measures of student academic achievement relevant in comparing the effectiveness of instruction when comparing different instructional delivery formats (online versus face-to-face) within a discipline of coursework, this study, then, builds on this body of literature using a quasi-experimental research design to compare measures of student success and student satisfaction in online and on-site courses offered in a College of Education at a large, urban university in the Northeastern United States.

The purpose of this study is to determine if there are any significant differences between the levels of achievement and satisfaction among students in online and on-site classrooms within courses offered under the specific discipline of education studies. This study acknowledges the need for and value of discipline-specific analysis when evaluating and comparing methods of instructional delivery. Across its specialized areas of research and practice, the study of education brings together theories, research, and practice based within psychology, history, philosophy, and sociology that reflects a unique and complex look at the reality of education and trying to understand how it is practiced. As it seeks to examine and describe education processes and systems, education studies, in its varied approaches and sub-disciplines, incorporates shared learning goals and outcomes. Education addresses shared content as it engages the content and development of curricula, teaching and learning, the relationship between ability, opportunity and success, and the policy issues impacting all of these (Bartlett & Burton, 2007). The purpose of this study's focus on courses offered within a single College of Education is to consider whether these shared learning goals and outcomes can be advanced equally when presented through varied instructional formats.

Research Questions

This study addresses several questions that relate to student performance and success in Education courses and how their success relates to perceptions of instructional quality as determined by student evaluations of teaching. First, it evaluates the academic achievement outcomes of students in face-to-face, hybrid, and online Education courses across three academic years (six semesters). As the purpose of this study is to consider if delivery method impacts student achievement, final grades earned are the metric used to evaluate student success. In this evaluation, additional consideration is given to academic preparedness of students, as indicated by entering GPA for new students and start of semester institutional GPA for continuing students. As an additional measure of academic success, this study evaluates rates of course completion between face-to-face, hybrid, and online Education courses to consider if delivery format impacts the likelihood of students completing the course. Student perceptions of course delivery and experience are discussed using SETs and perceived student satisfaction with Education courses offered in different formats is evaluated. Finally, student characteristics are examined to consider possible variations between face-to-face and online student populations and their potential impact on measures of student academic success and satisfaction.

RQ1: Is there a significant difference in student academic success, as measured by course grades, between Education courses offered through different instructional methods (online/hybrid/face-to-face)?

RQ2: Is there a significant difference in measures of course completion between Education courses offered through different instructional methods (face-to-face/hybrid/online)?

RQ3: Is there a significant difference in reported student satisfaction with either the course materials or instruction for Education courses offered

through different instructional methods (online/face-to-face) as indicated by a course level analysis of SFFs?

RQ4: Are there significant differences in student characteristics of sex, race/ethnicity, and residency between students who enroll in face-to-face as compared to online or hybrid courses?

Importance of the Research

As higher education institutions face difficult financial constraints and technology advances and becomes more ubiquitous, online course offerings and opportunities for online learning continue to grow. Online delivery of higher education appears to offer multiple affordances for HE institutions: using technology-enabled instructional tools may reduce administrative and instructional overhead and offer potential cost reductions; individual, flexible learning may translate into increased revenue generation by expanding the potential student body to include more non-traditional students and by re-engaging those students who have not been historically successful in traditional HE learning environments or who require more differentiated learning experiences; and online learning may allow institutions to broaden learning experiences by expanding the classroom with cross-cultural experiences that incorporate global perspectives relevant to the 21st century workforce.

As online courses become more commonplace for students across HE institutions, it is critical to consider the impact of online learning on student success in specific disciplines. The literature review in the current research describes how much of the current research in technology-enhanced learning has focused on specific learning tools and pedagogical practices. Some studies have looked broadly at distance education/online learning as compared to traditional classroom learning with no significant differences identified. Yet, more comparative research is needed to

understand the impact of the method of course delivery on student success and student satisfaction, as it relates to courses within specific subject areas and disciplines. Subject-specific research, as this project engages, is needed to better understand how course content intersects with the method of course delivery to provide administrators and instructors a more nuanced understanding of the potential impact of increasing online learning offerings within a discipline.

Summary

Responding to the increased demand and opportunity for online learning, providing benefits for both the institution and the student, administration and instructors need to understand the effect of different modes of delivery on student success. This study evaluates the effectiveness of online learning versus traditional courses offered over a three-year period within the discipline of education studies and examined the differences in student persistence, success, and satisfaction. The information in this introduction provided a justification for this study, as well as a brief overview of the literature and proposed research questions.

CHAPTER 2: LITERATURE REVIEW

Distance learning or distance education (DE) has been defined through association with broad and disparate instructional tools including early technologies such as correspondence courses, educational television and videoconferencing; through more modern Web-based applications like Learning Management Systems (LMS). The modern LMS, in turn, enables an ever-expanding collection of advanced web-enabled collaboration tools such as discussion boards, chats, electronic portfolios, blogs, and shared digital whiteboards; and delivers an ever-increasing complexity of multimedia and interactive content and personalized assessment and feedback. Studies of distance learning in its earliest iterations concluded that technologies were not significantly different from regular classroom learning. Given this, online education initiatives held greater appeal for Higher Education (HE) administrators on the basis of cost efficiency or the need to provide access to learners in settings where face-to-face instruction is not feasible. Instructors, meanwhile, could be assured that these DE courses offered increased learning opportunities that were essentially no worse than traditional classroom instruction in terms of student outcomes.

Yet, given the continual advances in instructional technology and the ever-increasing presence of DE across all types of HE institutions, the question of the relative efficacy of online and face-to-face instruction should be continually investigated, if, for no other reason than to reaffirm the parity of learning opportunities for both classroom and DE students. Moreover, interest in hybrid methods of instruction that blend in-class and online activities is increasing. As this review of the recent literature demonstrates, administrators, instructors, and policy-makers continue to seek a more nuanced understanding of the conditions under which online learning is most effective, as wells as explore the affordances and constraints for student success when engaging in different learning modes within different academic disciplines.

Research Focus on Pedagogical Tools and Best Practices

A number of studies have focused on specific pedagogical aspects of online course delivery. For example, Dalelio (2013) looked at the relationship between online discussion and student performance and categorized students based on their discussion board activities. He found active engagement in discussions to have a "positive relationship with student learning and performance" (p. 267) but did not address the type of content being learned. In another example, using a mix-methods approach including survey data and an analysis of journal entries, Smith, Sorensen, Gump, Heindel, Caris, and Martinez (2011) considered the role of group work in online learning, given the perceptions of students that online learning is a solitary and individualized process (p. 122). Considering the students preconceived notion, it is not surprising, perhaps, they concluded that "overcoming student resistance to group work" (p. 127) is more difficult in the online versus the face-to-face environment.

Focusing on social annotation tools (SA), Novak, Razzouk, and Johnson (2012) conducted a review of the current literature examining the pedagogical applications for technological tools used to "enable users to collaboratively underline, highlight, and annotate electronic text" (p. 40). Although their review of the current literature revealed a list of possible benefits from using these tools in instruction, like "participating and engagement . . . improve instruction . . . promote attention . . . improve reading comprehension … [and] peer-critique skills" (p. 40), they noted that in many instances these benefits "have not been supported by a rigorous empirical research" (p. 40). Out of more than 90 research papers identified in their study, they found only 16 experimental or quasi-experimental design studies that met their criteria for rigorous research to include in their discussion, of which they discussed in detail eight experimental and quasi-experimental studies specifically evaluating the instructional benefits and uses of SA.

Students using SA in these studies demonstrated improvements in critical thinking, meta-cognitive skills, and reading comprehension. Novak et al. also reported positive student response to the experience of using SA (p. 47). Yet, Novak et al. acknowledged the limitations of these findings as they noted there remains "little understanding under which conditions and within which contexts these tools should be implemented" (p. 49) and they advocated for further research.

In addition to looking at specific pedagogical tools, a portion of the literature remains focused on identifying generalized best practices for teaching at a distance. Keengwe and Kidd (2010) emphasized the role of the faculty and the barriers to online teaching in their review of the recent literature on best practices for online learning. They also identified the "unique attributes" of online learning "such as flexibility – anytime, anyplace – along with time for reflection and learners' anonymity" (para. 18). Discussing the work of numerous studies, Brinthaupt, Fisher, Gardner, Raffo, and Woodard (2011) summarized the key findings of best-practices-oriented DE research, including: "communicating clear goals and expectations; incorporating multiple active learning opportunities; providing frequent, prompt, and constructive feedback; and creating teacher support resources" [para 2.], all of which may seem applicable when delivering any kind of instruction either in a traditional classroom or through DE.

Studies focusing on the value and application of specific pedagogical tools and their applicability to the online learning environment offer some helpful first steps in building a corpus of best practices and instructional tools that enhance student learning. However, the lack of consideration for the differences introduced by specific learning goals (e.g., objective versus subjective knowledge, social literacy, or demonstrating applied learning) or subject areas (Math, engineering, comparative studies) leave open the question of what is the more appropriate or more desirable approach or tool to use when addressing specific student learning outcomes. Additionally, while focusing on

online learning/DE acknowledges the affordances of this delivery method versus

traditional classroom instruction, the resources and institutional dedication required in

overcoming the kind of barriers to increasing faculty adoption of DE that Keengwe and

Kidd (2010) pointed out presupposes evidence of *improved* – not merely equitable or

differentiated – student learning when deciding which mode of instruction to use.

Need for Comparative Evaluation of Online and Hybrid Learning

Some have suggested that online education will be the panacea for the ever-rising

costs of HE. Looking at three institutions: University of Maryland University College

(UMUC), University of California Online (UC Online), and Udacity (developed at

Stanford), Ben Wildavsky's "Classes for the Masses" (2013) examined one of the more

significant threats to the current delivery model of traditional HE course offerings – a

threat that speaks to both the financial pressures described by Bowen (2013) and

Kamaenetz (2013) and the impending technological disruption predicted by Christensen

(2011): low-cost online education. In reviewing the development of UMUC's very

successful program with over 120 offerings, Wildavsky pointed out that UMUC devoted

considerable resources to the design and development of their online courses. Early on,

UMUC recognized that "simply videotaping lectures is ineffective" (p. 128) and, instead,

built a system that ensures active faculty and student participation, detailed activity

tracking, and intensive faculty training (pp. 128-131); all of which required significant

investment of institutional time and resources. UC Online, which Wildvasky described in

counterpart to UMUC, has been struggling to establish itself, as it stumbled over faculty

resistance and issues with course credit sharing between campuses (p. 134). And, while

the promise of access to low-cost university education remains an attractive prospect

within the world of Massive Open Online Courses (MOOCs), Wildavsky noted that "the

cost of developing and delivering Udacity [an outgrowth of Stanford University] classes is substantial" (p. 143) and there is, as yet, no clear plan for long-term sustainability under the current open source model.

As Wildavsky's examples emphasize, quality online education, in its current form, requires a significant investment of time and resources, as well as faculty buy-in and campus support, to run effectively. Knowing where and how to invest those resources and being able to identify which programs will lead to the significant success of UMUC and which will become mired down in political fray of UC Online remains elusive without additional research. To be sure, online and hybrid learning has gained much popularity for its purported ability to deliver effective student learning at a substantially lower cost. Yet, Bowen, Chingos, Lack, and Nygren (2013) noted that "there is very little rigorous evidence on learning outcomes for students receiving instruction online" (p. 59). A recent report by the U.S. Department of Education (2010) analyzing studies of different delivery methods seemed to support the idea that online instruction could equal or improve upon traditional classroom instruction, but it too identified a need for further research to identify the particular strengths and weaknesses of online learning and the factors implicated for student success.

Considering the nature of this research project and the desire for more comparative evaluation to better inform the distribution of limited resources, a review of the current literature in Media Comparison Studies (MCS) seems appropriate. However, given the surprisingly vast corpus of research in this area, a select review of meta-analyses will serve to produce a summary view of the current findings and implications of this research. Indeed, many major studies have compared online to offline courses in experimental or quasi-experimental studies and found comparable student learning outcomes. In their criticism of "Old Concerns with New Distance Education Research," Lockee, Moore, and Burton (2001) expressed concerns about research that seeks to

determine which method is "best" as citing many early comparisons that seemed focused on defending one particular instructional method over the other (p. 62). Given this inherent bias in design and purpose, the researchers were not surprised, therefore, that most of these MCS studies resulted in "no significant difference" findings (p .62). However, they applauded the more current research trends that have placed greater emphasis on "content to be learned, the role of the learner, and the effectiveness of instructional design decisions" (p. 62) and they looked for more nuanced findings available in future studies.

Gunawardena and McIsaac (2004) conducted a critical analysis of research in DE and constructed a brief history of its development and components, in which they too pointed to the plethora of "early media comparison studies that yielded 'no significant differences'" (p. 387) and served primarily as institutional justification for developing or designing online programs. Bernard et al. (2009) cite a meta-analysis conducted by Russell (1999) "based on a collection of 355 comparative studies" (p. 1244) that found no significant difference between DE and traditional classroom instruction. Bernard et al. (2009) reviewed a second analysis from 2004 that looked at 232 studies conducted from 1985 to 2002 and found mixed results for DE course with more positive achievement and student attitudes, but poorer course completion rates. But, as Bernard et al. (2009) cautioned, because of the differences in "design, pedagogy, and technology," from these high-level analyses, little could, ultimately, be gleaned from the current research findings of "no significant difference" (p. 1244). Ultimately, they questioned whether additional research is even warranted given the inherent difficulties in separating instructional design from delivery method.

In 2009 (revised in 2010), the US Department of Education released a "Meta-Analysis and Review of Online Learning Studies" that looked at over a thousand empirical studies from 1996 through July 2008 to focus on studies that "contrasted an

online to face-to-face condition" and "measured student learning outcomes" (p. xiii).
Rather than "no significant difference," this comprehensive USDE analysis found that
"students in online conditions performed modestly better, on average, than those learning
the same material through traditional face-to-face instruction" (p. xiv). It was noted that
"studies in which analysts judged the curriculum and instruction to be identical or almost
identical in online and face-to-face conditions ha[ve] smaller effects" (p. xvi), which
implied a greater need for research looking at similar curriculum.

Indeed, a recent report from the National Bureau of Economic Research based on
a fully randomized study of online learning conducted by Figlio, Rush, and Yin (2010)
where students were randomly assigned to either an online or on-site format contradicted
the finding of the USDE and reported a greater level of student learning in the face-to-
face version of a course. Yet, like the USDE report that emphasized the need for
comparing similar curriculum when evaluating delivery method and student outcomes, in
this rare example of a fully experimental research design, Figlio et al. also worked to
ensure the online and face-to-face version of the course were identical in as many
characteristics as possible, such that the course was "taught by one instructor and for
which the ancillaries for the class, such as the web page, problem sets and TA support, as
well as the exams, were identical between the sections" (p. 5). As Figlio et al. asserted, if
additional comparisons of online versus on-site delivery methods using similar rigorous
research methodologies and design continued to confirm greater student success in on-
site or traditional teaching, then the rush to online education may be at the detriment of
students throughout HE. Despite the existing body of MCS research, they advocated for
more, not less, comparative research (ideally following a fully experimental design) to
better understand the impact of delivery method on student success (p. 21).

In addition to the need for MCS research that uses comparable course design and
random assignment, like that of Figlio et al., the need for research focused on particular

aspects of DE, such as instructional design strategies and subject-specific instruction remain. The complexities of learning, as well as the complexities of instructional design and delivery make it difficult to find external validity in these findings of "no significant difference" between online and face-to-face instruction within what is acknowledged to be a substantive body of MCS research. Rather than discouraging further studies, Gunawardena and McIsaac's (2004) summative history of MCS research pointed to the need for more research focused on major issues, including "effective instructional design, selection of appropriate media based on instructional needs, basic evaluation, and programmatic research" (p. 372). And they noted that "there appears to be little reported systematic research in this area because of the time and costs involved in conducting such large scale projects" (p. 372). At the same time, more students each year enroll in online learning with the expectation of achieving the same (or perhaps better!) learning outcomes as the same course delivered in a traditional classroom. If policy makers are to understand the value of allocating resources to growing online course offerings and the impact of those online offerings on student learning outcomes, research that addresses the concerns raised by the existing corpus of previous "no significant difference" findings and its variance found in experimental designs must be pursued. This project, while not adopting a fully experimental design, sought to narrow the focus to subject area (education) and, within this academic field, include some consideration of learner characteristics that might influence student enrollment into courses offered via different instructional methods.

Subject-Specific Research

A recent trend in online learning research considers the relationship of subject area and how online learning successfully navigates, or not, teaching the required

proficiencies in areas such as: sociology (Forsey, Low, & Glance 2013), service learning (McGorry 2012), counseling (Murdock, Williams, Becker, Bruce, & Young 2012), executive education (Rafferty & Anderson 2013), and history (Schrum & Sleeter 2013). Williams (2006) analyzed 25 such studies published from 1990-2003 comparing distance education and classroom instruction courses in the allied health professions and, using course grades as the primary measure of student achievement, and found a small but significant positive effect size for DE.   Acknowledging that course grades are a problematic measure of learning outcome achievement (Hu, 2005; Suskie, 2009), especially when there is no common examination (Bond, 2007), Williams also looked at a number of other instructional design elements in the research, such as "laboratory exercises, demonstration, group discussions, small-group projects, student presentations, guided imagery" and etc. (p. 130) as a measure of knowledge/skill building. Williams posited that the demonstrated improved outcomes for allied health students were, in part, due to the type of students who choose to study in the allied health professions, in that they were "usually older, highly motivated, and self-disciplined" (p. 135), all of which were known student characteristics that Williams noted the research had already showed to be positively correlated with successful online learning.

Similar to the experimental design approach of Figlio et al. (2010), Dutton and Dutton (2005) conducted a comparison of a business statistics course offered with identical content and assignments in either online or face-to-face format.   As with their previous study comparing computer programming classes (Dutton et al., 2001), this study focused on the strength of the parity between the sections including shared readings, assignments, and assessments. A pre-post survey was administered to evaluate the differences in student attitudes and motivations. While significant differences were found in the academic performance of the students, with the online students performing at a much higher level than the face-to-face students, Dutton and Dutton attributed this

difference to other factors, including factors that appeared to be the motivating factors for students in the self-selection process, such as age, entering GPA, and working hours outside of academic study.

Likewise, Margoniner (2014), an instructor, also compared two versions of a single course offered in different formats: his introductory astronomy class for general education, which he taught as both a "flipped face-to-face" and an online course (p. 298), both of which cover identical content. The "flipped face-to-face" version devoted class time for "discussion and synthesis" supplemented by a textbook, online pre-test, and a "set of video mini-lectures" to be viewed before each class (p. 298). In the online version, the in-class activities and discussion were replaced with discussion forums that are evaluated using a rubric. Although Margoniner found that "the online student population is significantly different from the one in the flipped class" (p. 300) with, as found by Williams (2006), the online students tending to be older and/or more advanced in their degree programs, she also found "no significant difference" between the learning outcomes of students in either course delivery method and noted significant learning, based on pre-test/post-test results, for students in both modalities. As with Williams' study, Margoniner noted that students who elected to take the online course may have done so because they had the characteristics of a more successful online learner. However, in this case those characteristics were not related to the subject area of the course – which was a general education course serving students from across the disciplines – but as a result of the self-selection of the students themselves.

Also referred to as Media Comparison Studies (MCS), limited research exists comparing online learning and face-to-face instruction within the discipline of education studies or teacher preparation programs. The 2009 USDE meta-analysis of online learning studies found only 10 studies related to teacher education. Due to critical teacher shortages, Special Education has recently emerged as a leader in the adoption of

online instruction to reach teacher candidates in rural and otherwise underserved areas. This has resulted in a greater number of MCS focused on the specific instructional needs and outcomes for special education teacher preparation programs.

Keramidas (2012) examined a particular behavior management course "for students enrolled in the early childhood/early childhood special education dual certification program" at a "large university in the Mid-Atlantic United States located in a highly rural area" (p. 27). Half of the students completed the course as a traditional face-to-face course. The other half "viewed archives of the face-to-face courses on their own time" (p. 27). Other than how lectures were delivered, the requirements of the course were identical in both the online and face-to-face sections. At the end of the term, the online section exhibited higher attendance (with "two sessions being the highest amount of sessions any one student missed" (p. 28)); lower overall mean on quizzes, assignments, and the midterm exam; and lower overall grades in the course. Keramidas suggests that student expectations of a less rigorous course, as well as some struggles with the technical aspects of managing the online course could have influenced the students' success rates in the online section (p. 31). While this study provides a useful example of a comparison of an education-related course offered in both online and face-to-face format, Keramidas does not consider whether a course which merely records and redelivers lectures electronically acceptably represents an online course that has been fully adapted to incorporate the pedagogical best practices for online instruction (such as opportunity for classroom discussion and student interaction).

Thompson, Klass, and Fulk (2014) focused their research on a specific introductory course in special education offered at a large Midwestern public university. In their review of the literature, they note "no recent studies of students' achievement and satisfaction following online versus f2f [face-to-face] instruction" (p. 230) in the field of special education teacher preparation. The course in their study consisted of identical

content and requirements except that the students in the face-to-face section attended a large lecture during specific days and times each week and the online students could access the same lectures through enhanced podcasts that could be accessed on-demand (p. 231). Notably, while the students self-selected into either the online or face-to-face sections, unlike in other studies that pointed to younger students enrolling in the traditional classroom course, the online section in this study "had proportionately more freshmen (34% vs. 12%) and fewer juniors and seniors (16% vs. 41%) than the face-to-face section. Overall, the researchers found that online students performed slightly lower than face-to-face students but reported slightly greater levels of satisfaction with the course and equal levels of engagement. When student characteristics such as gender and attendance were considered there were, once again, no significant differences between achievement measures in either section. As a result of their research, Thompson et al. suggested that the difference between online and face-to-face should present little barrier for a content-heavy course such as an introductory special education course and recommended that further research focus on instructional components and methods for improving instructional technique and maintaining quality in the online environment.

Moving beyond the introductory courses that focus primarily on transmitting objective knowledge rather than developing skills and professional practices, Rakap, Jones, and Emery (2014) evaluated the success of a four-course professional development program that followed a traditional special education program of study in autism and related disabilities. Despite participants reporting "themselves to be more knowledgeable and competent in all areas targeted by the project on completion" (p. 13), the researchers noted that participants subsequently reported infrequent application of the teaching strategies learned in the program. They noted these results were consistent with prior research "investigating the effectiveness of web-based courses and programs helping teachers improve their knowledge and skills about working with children with disabilities

and their families" (p. 13) and indicated the difficulties inherent in adapting new instructional practices to classroom contexts may pose a barrier to teacher professional development programs, regardless of the instructional delivery method (p. 14). As this study was based on a pre-test/post-test design rather than a stricter comparison of online programming with traditional classroom instruction, a more comprehensive study was suggested. However, it was noted that significant learning outcomes and student satisfaction were identified in the web-based courses as offered.

Keengwe and Kang (2012) conducted a meta-analysis of research examining blended learning in teacher preparation programs, where blended learning "includes a course that combines online and face to face (f2f) learning" (p. 82). As they argued, teacher education includes some unique elements, such as "field experiences, student teaching, and so on" (p. 64) that warrants separate consideration from other higher education programs and poses some significant challenges to fully online delivery. In their analysis of the literature covering tools used in integrating online and face-to-face learning, the researchers found that technology tools, like discussion boards, could allow student teachers to confer in a "more open community" and technology tools, in general, provided the opportunity for student teachers to integrate "creative ideas into practical skills and practice" by allowing for "real and authentic experiences" (p. 86) (teachers used technology to prepare learning materials to be used in their classrooms). They also noted the value of social engagement using online tools like Ning, as cited in one study by Arnold and Paulus (2010). However, the research they reviewed also pointed to barriers for students who "were struggling with the use of new technology tools, their necessity, and effectiveness" (p. 86). They conclude their review of the research noting the need for curriculum designers to "integrate traditional conceptual frameworks into online activities" (p. 89); for "more empirical research . . . to examine the effectiveness of blending learning in TPPs [Teacher Preparation Programs]" (p. 90); and for greater focus

on "theoretical and practical understandings" of online environments for teachers who "will face these issues when they go into their own classrooms" (p. 90).

As this brief overview has shown, Special Education, focused on the pedagogical needs of students with disabilities, is one aspect of educational studies that has received more focused attention by researchers evaluating the effectiveness of online learning. Vernon-Dotson, Floyd, Dukes, and Darling (2014) reviewed 17 studies published between 2000 and 2012 in their investigation of course delivery methods for preparing special education teachers. They noted that most existing research focused on broad categories of teacher education programming at the macro level or course analysis at the micro level with little research focused on the in-between. They chose, therefore, to focus on course delivery in the specific area of special education teacher education to consider changes in course delivery over the past decade, "global factors [that] contribute to the effective course delivery" (p. 37), and other benefits that support one method of course delivery over another in this specific subject area. Vernon-Dotson et al. (2014) found that online learning met a particular need for special education programs in "rural and remote areas where colleges and universities are not readily accessible" (p. 41) and where shortages of qualified teachers are notable. Both convenience and the ability to reach a wider student base was noted as a benefit of online learning found in the research they reviewed. And, once again, positive or "no significant difference" was found between DE and traditional classroom instruction modes of delivery for special education courses in studies "where participants were able to self-select the preferred mode of instruction" (p. 41). Critical factors they identified for successful online learning in special education courses included similar suggested best practices for all online learning environments – as previously discussed and summarized by Brinthaupt et al. (2011) – such as sustained and substantive opportunity for online discussion, a quality instructor, and "a standardized evaluation system for distance education" (p. 42). Also noted was the

need for sufficient infrastructure and support to avoid or quickly address any technical difficulties experienced by the students (p. 44). As with the other research, this suggestion did not seem specific to the needs/goals of special education learners as compared with any other online learner.

In light of the fact that studies focused on specific courses or narrow sets of courses within specializations have, when controlled for different characteristics of the students in each modality, also yielded either "no significant differences" or, at best, mixed results, the value of additional research at the specific course level remains unclear. However, the value of a body of research at the discipline-level, as reviewed by Williams (2006) for allied health, may provide a clearer understanding of the type of student who would best succeed in educational studies courses offered in a particular instructional method. Additionally, in focusing on a discipline within a college, this study speaks to the research gap identified by Vernon-Dotson et al. (2014) as it is situated in-between the range of studies previously conducted at the all-encompassing macro-level studies at the broad institutional level and the individual course comparison.

Use of Grades to Evaluate Student Success

This study uses course grades to compare the level of learning between different instructional methods. Course grades are typically assigned by individual instructors at the end of most courses and universal agreement between instructors on what constitutes an A, A-, B+, and so on remains elusive. The very independent and individualistic nature of this assessment, along with the known variation in systems of grading across institutions, disciplines, and courses has fueled criticism as to the arbitrary nature of this evaluation of student learning. Critics, such as Clayson (2009) have questioned the validity and reliability of using course grades for the purpose of assessment and

evaluation of learning. At the same time, "grading is one of the fundamental facets of education" (Schneider & Hutt, 2014, para. 1) and it is has been used repeatedly as an accepted measure of student learning at multiple levels, including the student, department, and institution (Burns, 2011).

This study accepts the findings of previous research, as detailed and supported by the work of Canfield et al. (2015), which reviewed the grades from 2,400 undergraduate courses, including 5,000 to 6,000 students, over a thirteen-year period. Of particular note for the purposes of this study, their comprehensive analysis affirmed a strong "trait relationship" (p. 10) whereby grades within a specific discipline were shown to correlate more strongly than grades between disciplines. Canfield et al. thereby affirmed the assumption of this study that grades may be considered a valid measure of learning within a discipline and that, within that discipline, separate course grades may be combined as a valid measure of the discipline. This assumption is further supported by the fact that this study involves data from a single institution, which further reduces the factors of variance noted by Canfield et al.

Using Course Completion as a Comparative Measure Between Instructional Methods

Cochran, Campbell, Baker, and Leeds (2014) identified retention for online classes as a "continued issue" with notably higher attrition rates for online courses than for face-to-face classes. Other studies have affirmed this concern regarding persistence in online courses (Allen & Seaman, 2015; Patterson & McFadden, 2009). A review of 40 studies published over an eight-year period by Muljana and Tian (2019) supported the assertion that online learning resulted in lower student retention, if a broad set of "factors influencing online student retention" (p. 25) were not adequately addressed. Of import for

this study, Muljana and Tina noted the need for carefully defining the parameters for measuring retention before conducting comparison between instructional methods. Additionally, both Cochran et al. and Muljana and Tina identified a series of student characteristics that were shown to influence rates of online course completion.

While some studies have addressed this issue for business programs (Johnson-Lutz et al., 2015), limited research has focused specifically on comparisons of course completion rates within the discipline of educational studies. This study expands on this aspect of the research to consider the relationship between rate of course completion and instructional delivery method for courses within a College of Education.

Using Student Evaluations to Assess Learning

In its early development, online learning often appealed to students who were not well served by traditional learning institutions and offerings. In the 2002 report *Growth in Distance Education* the GAO reported that online students, as compared to their traditional counterparts, tended to be older, to work full-time, to study part-time, and to be married. Given the growing number of current students likely to take at least one online course during their academic program, it is not surprising that Blackmon and Major's (2012) review of the qualitative research on student experiences with online research acknowledges that students who take online courses are now "in many ways the same students who take courses offline" (p. 77) and point to a need to look more closely at specific populations and aspects of online learning in order to better understand what does and does not influence student success. When Beck and Milligan (2013) examined the correlation of generally accepted predictors of academic success and student success in online programs, they found an unexpected absence of a significant relationship, and a greater impact from the academic and social environment of the students. Their research

suggests that even well-researched findings using learner characteristics known to influence student success in face-to-face higher education programs may not fully encapsulate the factors surrounding student' success in online coursework.

Rather than scrutinizing specific student characteristics, this project used a University-developed instrument: student feedback forms (SFFs) to evaluate student satisfaction. There exists some controversy around using student evaluations of teaching as a secondary indicator to assess levels of student learning. While it may be assumed that student feedback directly speaks to student opinions of the course (satisfaction), it is of interest to briefly consider the literature around the validity of the relationship betwen student student evaluations of teaching and student learning.

Building upon two meta-analyses conducted by Cohen (1981) and Clayson (2009), Stehle, Spinath, and Kadmon (2012) investigated the relationship between Student Evaluations of Teaching (SETs) and quality of teaching. Citing a number of studies, Stehle et al. assert that: "the most apparent criterion for teaching effectiveness is student learning (e.g. Cohen 1981; Marsh and Roche 2000; Murray 2005; Svinivki and McKeachie 2010)" (p. 889). At the same time, Stehle et al. agree with Cohen (1981) who understood that there could be no criterion that would exactly correlate with teaching effectiveness, as performance on a final exam, for example, could be impacted by other factors outside of teaching effectiveness (student health, student attitudes, environmental factors, etc.) and therefore the researcher must identify a criteria that "'show at least a moderately strong relationship to this index' (p. 281)" (p. 890). In an effort to speak to the debate around the validity of using SETs to measure teaching effectiveness, Stehle et al. reviewed studies that have attempted to "compare SETs with criteria of student learning" (p. 889). Despite several substantive studies, they found that there exists on-going controversy around using SETs as evidence of teaching effectiveness, as critics argue that SETs reflect more about the student (e.g., interest in the subject, personal

taste), the class configuration (e.g., class size, workload), and the personal attributes of the instructor (e.g., personality, lenient grading) rather than the effectiveness of teaching.

Cohen (1981) conducted a meta-analysis of "41 independent multi-section validity studies" (p. 890), which revealed relatively strong correlations between the criteria for student learning to SETs, but found some key areas where inconsistencies were more prevalent. Cohen identified several factors that accounted for about a third of this variability: including where "instructors were graduate students . . . SETs were collected after the students had learned about their final grades and . . . [where] the exam was graded by the instructors themselves" (p. 890). He further identified some difference attributable to whether the criterion was "an objective test or a non-objective test" (p. 890); however, Cohen determined that "this feature did not significantly predict the size of the validity coefficient" (p. 890). Therefore, except for these aspects of controversy, Cohen concluded that student SETs could be used a reliable measure of student learning.

Clayson (2009) conducted a subsequent meta-analysis including more recent studies and found older studies (before 1990) "yielding stronger correlations than more recent ones" (p. 891), as well as a notable relationship between the "journal in which the study was published" whereby "studies published in educational and psychological journals" reported higher correlations between SETs and student performance (p. 891). This finding reinforces the importance of examining courses with a shared or similar curriculum.

Stehle, Spinath, and Kadmon (2012) further examined a particular aspect of Clayson's findings as related to the "degrees of objectivity of the learning criterion" where "a weaker association with SETs" was found with "the more psychometrically objective" measure for student learning (p. 891). While Clayson concluded that students were less able to evaluate their own objective learning, Stehle et al. noted that "Clayson's results are not without ambiguity" due to a weakness in the way Clayson categorized

objectivity and the inclusion of "very different disciplines" (p. 891), which may have resulted in conflating subject matter with objectivity. Stehle et al. suggest that "these and other questions need to be addressed before accepting the conclusion that objective measures of learning are not or only weakly associated with SETs" (p. 891). Considering a more nuanced categorization of measures of student learning to make distinctions between the types of learning required to succeed on multiple choice exams versus essays versus practical exams, Stehle et al. posit that the "strength of association between SETs and students' objective learning varies with the kind of learning criterion" (p. 891). This, again, reinforces the need for the kind of subject-specific analyses proposed by this study.

In their study, Stehle et al. (2012) point to the work by Prosser and Trigwell (1991) which considered both an assessment of the quantity of student learning (through a more psychometrically objective exam) and the quality of student learning (as measured by an essay exam). Prosser and Trigwell found a positive correlation between SETs and performance on the measure of the quality of learning, but a less significant and slightly negative correlation between SETs and performance on the measure of quantity of learning. As Stehle et al. interpret these results, measures of quantity of learning may be more greatly impacted by factors outside the classroom (e.g., studying the textbook to acquire factual knowledge), whereas measures of the quality of learning, which require more analysis and explanation by the student, may be more strongly correlated to teaching effectiveness and therefore SETs .

The purpose, therefore, of Stehle et al.'s (2012) own study was to further expand on Prosser and Trigwell by examining the relationship of SETs on both achievement on a practical examination and a multiple-choice test. In summary, their study included five sections each semester of a two-week intensive module on visceral surgery with 30 students in each section and a different instructor for each section for a total of 839 students (500 female) and 21 unique instructors. Measures used included a measure of

students' prior academic achievement (four grades from previous courses calculated into a compound score); and the SET, which included 15 items, was used to create an Overall Instructor Quality score (a combination of 4 items), an Overall Course Quality score (single measure), and a Students' Subjective Learning score (single measure). Examination scores were collected from the multiple-choice test and an Objective Structured Clinical Examination [OSCE]. In order to protect validity, the same multiple-choice test was not used each semester and the total possible points varied in different terms, so the researchers regularized the scores for analysis. The OSCE utilized competence-based stations with simulated patients evaluated by an impartial examiner (not instructor) using checklists and scores were, again, regularized for the purpose of analysis. Additionally, in the preliminary correlations, it was determined that sections "were not equivalent with regards to prior academic achievement" (p.895) and, as a result, this variable needed to be held constant when conducting further analysis.

Analysis revealed that Overall Instructor Quality, Overall Course Quality, and Students' Subjective Learning "were all strongly positively correlated (p. 896) with performance on the OSCE, whereas none significantly correlated with multiple-choice score. Data thus supported both of Stehle et al.'s (2012) hypotheses, suggests "that there is a common factor driving the correlation between SETs and OSCE-scores, but that this common factor is missing with regard to SETs and multiple-choice test scores" (p. 897). Student's performance on the multiple-choice test also correlated strongly with their previous academic performance. Stehle et al. posit that this might be because the course grades used to calculate the previous academic performance may have been largely based on similar multiple-choice tests. However, it may also be the case that the material covered in the multiple-choice test closely related to the material covered in the previous courses.

In what may be understood to be a response to Cohen's concerns about selecting a criterion that adequately assesses student learning outcomes, Stehle et al. chose to investigate both a highly prescriptive program (medical training) and two measures of student learning: a multiple-choice test and an Objective Structured Clinical Examination [OSCE] that correspond highly to a set of specific and demonstrable learning outcomes for the students. Stehle et al. conclude that the correlation between SETs and the OSCE validate their assumption that "more effective teaching leads to both superior student evaluations" (p. 899) and better student learning outcomes on practical examinations. Since, they argue, multiple choice tests rely primarily on testing knowledge that can be gained through close examination of the textbook or other independent study, these measures of student learning may be less influenced by the quality of the instruction provided (as students have methods to compensate for poorer quality instruction). Additionally, they contend that instructional focus on high-level thinking and analysis throughout the course may not coincide with the cognition and skills required to successfully complete a multiple choice exam. This would be another reason why student perceptions of teaching effectiveness correlate more directly with practical examination. Further, they suggest that since multiple choice tests "are the most frequently used exam format" in the current higher education system, this may account for why Clayson (2009) did not find a positive association between SETs and student learning after 2009 (p. 899).

Research supports the correlation between student feedback and performance in fields that require practical examinations and/or demonstrations of applied skills and given. This project focused on courses offered in education studies, which – while not exclusory of multiple-choice or other objective assessments of learning – in the majority, focus on developing practical knowledge and building performance skills in such areas as leadership, social justice, diversity, and classroom and conflict management. Thus, it

may be assumed that student evaluations of education courses, like those in the medical field studied by Stehle et al. (2012), exhibit a higher-level correlation between teacher evaluations and student learning.

Bias in Students' Evaluation of Teaching: Instructor Age/Sex

Peter Schmidt (2017) reported on a recent case brought against the University of Kansas by a faculty who was denied tenure. According to the faculty, Catherine A. Joritz, the university used anonymous student evaluations of her teaching as an important factor in considering her tenure case. Joritz, in turn, pointed to comments left by students that expressed what she perceived as underlying student bias against her cultural background. Schmidt notes that institutions have little firm legal guidance on how to appropriately use student feedback, even as the expansion of anonymous, online feedback systems has reduced potential risks for individual students who leave inflammatory or clearly biased comments. As the Kansas Board of Regents emphasized in their response to Joritz, the increased use of anonymous ratings by students corresponds to an even greater need to identify tools "corrected for major sources of bias as demonstrated by research" (p. A26). Similar to the issues brought forth by the problems of platform moderation on massive tools like Facebook, Twitter, and YouTube, HE institutions need to create clear policies that support honest and appropriate feedback, while understanding the complexities of anonymous online sharing. With the acknowledgement that this study did not seek to address the validity of the relationship between student feedback and instructor performance, the following is a brief summary of the research related to understanding the biasing factors inherent in SETs.

It is precisely because of the complexity of these issues that Jones, Gaffney-Rhys, and Jones (2014) cautioned administrators and faculty managers against using student

evaluations of teaching as a summative performance measure in high-stake decisions, like tenure decisions. They cite Pounder's (2008) examination of the previous literature relating to SETs, which considered the validity and reliability of student feedback as influenced by three broad groups of factors: student-related factors, course-related factors and teaching-related factors. With differing degrees, the literature addresses the ways in which each of these factor groups may impact the validity and reliability of SET instruments.

Teacher-related factors include sex[ii], age/experience, behavioral traits and teacher-influencing factors (including grade inflation). However, sex and age/experience cuts across both student- and teacher-related factors and represents a co-influencing factor (student on teacher/teacher on student). Wilson, Beyer and Monteiro (2014) cited several studies from the early 80's indicating a student preference for gender normative roles in their professors: female professors perceived as "warmer and more caring" received higher evaluations, while student evaluations of male professors correlated more strongly with perceptions of competence. Wilson et al. pointed to Bachen et al. (1999) that indicated that female students gave especially high ratings to female academics; and to Tatro (1995), who found that female students generally tended to rate their teachers more highly than did their male peers.

Noting that studies focused on sex or gender have resulted in mixed results, Centra and Gaubatz (2000) looked at the inter-relationship between the gender of the student and the gender of the instructor. Unlike previous studies that included simulations, single institutions, and small data sets, the researchers used a large number

---

[ii] The research, as discussed here, often conflates, or does not explicitly acknowledge the difference, when referencing the concepts of biological "sex" or socially-constructed "gender." As a result, the terms are used in alignment with how they were used by the authors of the studies discussed.

of different types of institutions and course subject areas and found that "female instructors received higher ratings on six of the eight variables when rated by female students. . . [while] Male instructors were not rated significantly different by female or male students" (26). However, their analysis across academic disciplines was more varied and not significant for courses in Education, in particular. Nor did their identification of some small bias in favor of female instructors by female students consider course delivery/format as a complicating factor.

Kogan, Schoenfield-Tacher, and Hellyer (2010) focused on both the gender and the age (tacitly inferred through position and rank, which is assumed to correspond with relative age) of the instructor. However, their study considered these factors and their potential implication for how instructors receive and respond to student feedback. They did not address the possible influence of these factors on the ratings assigned by the students. In their own study (2014), Wilson, Beyer, and Monteiro found that the age of the instructor was an influencing factor - sometimes positive and sometimes negative – in the student expectations and evaluations of their professors. The researchers based their findings on a relatively simplistic study using pictures of potential instructors, instead of analyzing data from existing student feedback of actual instructors. They indicated that additional studies should consider the age of the professor as a potential moderator of student evaluations of teaching.

In their preface to a discussion of how to mitigate its effects on SETs, Peterson et al. (2019) succinctly summarize the current research on instructor sex and SETs: the belief "that SETS contain biases against female instructors is widespread" (p.1), even though, a review of the research does not support such a definitive conclusion. Further research is needed to investigate the complexities of the relationship between characteristics of the instructor and student feedback. However, as detailed instructor

demographic data are not within the scope of this study, this aspect of the research will not be addressed here.

Summary and Approach to the Research

As this review of the current research reveals, the potential exists for technology to increase and enhance learning opportunities, and perhaps reduce costs and improve access. Despite the body of comparative research resulting in "no significant difference" between distance education (DE) and classroom instruction modes of delivery, more research is needed to identify and evaluate the best tools and pathways that result in student success within specific fields of study that are focused on a shared set of student learning outcomes. As the broad sweep of research examples presented here imply, the relative success or failure of any one strategy greatly depends on a complex interrelationship between the culture of the institution, the resources available, the technical skills of both the instructor and students, the course and student learning goals, and an understanding of the capabilities and limitations of the particular software or tool. This would seem to support additional research that reduces these differences through a focus on a specific institution, as well as a specific field of study, such as was conducted in this study.

Clayton Christensen (2011) characterizes current education research as focused on correlation instead of causation to the detriment of identifying and building a body of prescriptive approaches to teaching and learning that work for specific types of learners in specific types of environments (pp. 187-190). Instead, he suggests that researchers should be focusing on the "*anomalies . , .* [for] it is only when we find something that the preliminary correlation *cannot* explain that the research can improve" (pp. 190-191). Whereas current research practices have focused on *reliability* (p. 197) and *internal*

validity (p. 198), Christensen argues that a renewed focus on *external validity* (p. 198) would provide policymakers, administrators and reformers tools needed to identity the "recommended actions would be effective" (p. 199) for a particular student in a particular social and learning environment with a particular set of intelligences and aptitudes.

This research project sought to add to Christensen's (2011) call for educational research that is actionable and generalizable, while focusing on a specific subject field of courses: Education. A wide sampling from a large, urban institution in the Northeast U.S. was included in the study over a period of six semesters to facilitate the identification of trends and patterns within this field and in consideration of the characteristics of this institution and the students it serves. In addition, to improve the possible applicability of the results, the narrowed, but the longitudinal focus on a College of Education at a single institution that was used in this study accounted for some of the weakness in the correlation between student feedback and learning outcomes, as identified by Clayson (2009) that can occur when analyzing feedback from across disciplines or institutions. This focus on a field that includes an emphasis on practical and applicable skill building also takes into account the research design of Stehle et al. (2012), which revealed a stronger correlation between student feedback and learning outcomes for practical versus objective assessments. Additionally, with the narrow focus on a single institution and discipline, this study avoids some of the potential criticisms when using course grades as a measure of student success, as described by Canfield et al. (2015). However, as this project focused on a single institution, one limitation of the study is the need to replicate the analysis across a range of institutions by type, size, and geographic location, among other variables.

CHAPTER 3: METHODOLOGY

This study expands upon existing research addressing concerns around student success, student persistence, and student perceptions of learning in light of the growing trend for increased distance education courses delivered by more institutions and enrolled by more students from across the HE landscape. While online learning affords a range of potential benefits for both the institution and the student, policy-makers, administrators and instructors need to better understand the relationship between different modes of instruction and student learning outcomes and identify any potential barriers that prevent students from enrolling and fully participating in hybrid and online learning. As discussed in the review of the literature, many previous studies described issues with how to teach an online class or identified the specific challenges in developing online learning. This study builds on the work of Legutko (2007); McCrory, Putnam, and Jansen (2008); and Kirtman (2009); among others that explore the relationship between student success and satisfaction and course delivery format to focus on courses offered specifically within the discipline of education studies. Unlike other studies that focus on a few specific courses or specific degree programs within the discipline (e.g. master's programs or teacher preparation) to observe the differences in student persistence, success, and satisfaction, this study evaluated the effectiveness of hybrid and online learning courses as compared to traditional classroom teaching across all course offerings over a three-year period within a College of Education at large, urban, public University.

Research Setting and Participants

This study considered a data set from a single institution of higher education: a public, 4-year, urban institution with a student population of over 36,000 located in the Northeastern United States. According to the 2019-2020 enrollment period fact book

published online, overall, the University enrolls about 66% female students and 44% male (p. 37). The University particularly highlights its diverse student population (p. 12), which includes a 54% white, non-Hispanic; 12% African American; 12% Asian and 7% Hispanic population (p. 37).

The units for analysis for this study consisted of courses taught within the University's College of Education. Established in 1884, the College of Education was one of the first colleges in the institution and, in 2019-2020, it enrolled approximately 1,800 students, including 1,000 undergraduates and 800 master's and doctoral students (p. 42). The College of Education includes graduate and undergraduate degree offerings in varied areas of specialized study within educational studies, including, among others: adult and organizational development, community engagement, counseling psychology, higher education studies, human development, psychological studies in education, school leadership, and teacher preparation. Courses are offered through one of four possible designated course departments: Policy, Organizational, and Leadership Studies (POLS); Psychological Studies in Education (PSE); Teaching & Learning (T&L); or the "College of Education (CoEd)." The "College of Education" departmental designation is used where the course is 1) routinely taught by instructors from different departments; 2) required by cross-departmental degree programs; and/or 3) part of the University's general education offerings that are open to all students from majors across the University. Most courses in the College are offered through the department of Teaching & Learning, which, as expected, also includes the most majors. Course section enrollments in the College may range from a single student to over 80 students in larger lecture courses.

Course Scheduling and Delivery Format

Courses in the College of Education are offered in the traditional U.S. semester-based system. Students typically attend for two semesters each year: fall and spring. Each fall and spring semester includes 15-weeks of instruction, not counting breaks, study days, and exam periods. The University does offer two summer semesters: Summer I and Summer II, respectively, with each divided into two six-week accelerated terms. Additionally, some courses may be offered across both summer terms for a longer 12-week instructional period that emulates the traditional fall and spring schedule. Except in some specialized degree programs or where students stop-out and need to catch-up for on-time degree completion, summer coursework is not required. Students are considered as continually enrolled as long as they enroll in both fall and spring semesters until degree completion.

For any given semester, each course section is scheduled by the College scheduling administrative team within the academic operations unit of the College. Course delivery format is determined at the time of scheduling, which may be several months prior to the course start date. Instructional format for any given course section may be offered in a one of a range of available delivery formats including traditional, in-person, lectures and seminars; synchronous and asynchronous online courses; internship, practicum and field-based observation and participation courses; as well as independent study and graduate research opportunities. The University assigns each course an Instructional Method (in parallel with a designated Meeting Type) that can be used to assign a level of online learning representing the extent of online learning included in a given course as continuum from "entirely face-to-face" to "entirely online" with a number of variations in between. Table 3.1 shows the range of possible Instructional

Table 3.1	
Level of Online Learning	
Instructional Method	Level of "Online"
Online	13
Online with required synchronous meetings	12
Online with some face-to-face	11
Video conference	10
Classroom	9
Classroom & Applied Experience	8
Recitation	7
Studio	6
Laboratory	5
Learning Community	4
Community Based Learning	3
Non-Classroom Field Experience	2
Supervised Experience	1

Methods and the corresponding assigned "level" of online learning. As the number increases, the extent to which the course is online increases.

The Instructional Method for a particular course section is determined by a range of factors, including, among others: tradition or historical record of the course's offerings; availability of qualified instructors; instructor preference and/or background and technological comfort level; perceived complexity of the course content or learning outcomes; role of the course within a larger degree program; role of the course within a program of study approved by a state certification bureau or other accrediting body; student demand and enrollment trends; perceived need for flexibility in course scheduling by the administration; and availability of on-campus meeting locations. The final selection of the delivery format is determined through collaboration between the instructor, the course or program coordinator (if one is designated), the department chair, and the College's administrative scheduling team.

Students should not be considered as either randomly assigned or fully self-selected into a given course or course delivery format. As a result of the numerous variables that determine course delivery format for a given course section, students may elect to enroll in an online course because it is the format of choice or because it is the only format available during a particular semester that meets their degree requirements. A student wishing to avoid stop-out and to graduate within the projected degree completion period determined at the time of admission, may elect to enroll in a traditional face-to-face course when an online course is more desirable or convenient for that student or vice-versa. Thus, enrollment alone does not indicate a clear preference for online learning.

Student Feedback Forms for Courses

Registered students in courses across the institution are given the opportunity to provide feedback for each course using the University's e-SFF (electronic Student Feedback Form) system. Using this system, feedback forms are generated automatically for all courses with at least five enrolled students. Individual instructors may request forms also be generated for courses with fewer students. Some professional schools, such as Law and Medicine, do not use the e-SFF system for collecting student feedback. For those courses included in the e-SFF distribution, except for recitations and courses with more than one designated primary instructor, all feedback forms include the same standard 11 questions listed in Table 3.2. For any given course section, the instructor, course coordinator, department chair, or school/college administration may elect to add additional custom questions to the SFF, but the standard questions cannot be removed or omitted from the form.

Table 3.2

Questions on the Student Feedback Form

Q1: I came well prepared for class
Q2: The instructor clearly explained the educational objectives of the course
Q3: The instructor was well organized and prepared for class
Q4: The instructor was conscientious in meeting class and office hour responsibilities
Q5: The instructor promoted a classroom atmosphere in which I felt free to ask questions
Q6: The instructor provided useful feedback about exams, projects, and assignments
Q7: So far the instructor has applied grading policies fairly
Q8: The instructor taught this course well
Q9: The course content was consistent with the educational objectives of the course
Q10: The course increased my ability to analyze and critically evaluate ideas, arguments and points of view
Q11: I learned a great deal in this course

Students evaluate each of the SFF questions on a five-point Likert scale, including: "Strongly Agree (5)," "Agree (4)," "Neutral (3)," "Disagree (2)," or "Strongly Disagree (1)". Students may skip any question within the form and still submit the form. Skipped questions are recorded as "Not Answered (0)." For each course section, the University reports the total number of completed evaluations and the total number of registered students, along with the total number answered and the calculated mean of the scale for each question.

Each fall and spring semester, students are invited to complete and submit a feedback form for each registered course – assuming the course is within the University's distribution guidelines – using the University's e-SFF online system. Feedback is solicited beginning in the 12th week of instruction through the end of the 15th week of instruction before the final examination period begins and before final grades must be submitted by instructors. Completion of feedback forms is voluntary and confidential. Students are reminded to complete their feedback forms through a series of communications at both the University and College level. Programs and individual instructors may also encourage students to submit feedback. A student who completes a

form for every course in which he or she is registered in a given term receives access to

the SFF Data for Students portal that provides the opportunity for students to view

selected course and instructor data from prior semesters. It is assumed that students value

access to this portal, and are, therefore, further incentivized to complete feedback forms,

as they may then use this access to the student portal to influence their future selection of

courses and instructors based on the evaluations of other, former students. The

University does not publish historical or trend completion rates for the SFF forms,

although the response rate for the current term is published to the students throughout the

feedback collection period. At the time of this study, the University claimed that 43% of

feedback forms had been completed for the most-recent semester (2020).

Data Collection Procedures

Data that were collected from the University's Enterprise Resource Planning

(ERP) system – Banner – included course, registration, and student information. Reports

of the courses offered by the College of Education included in this study were drawn

from the available student, course and instructor records and registration data for each

spring and fall semester beginning Spring 2017 through Fall 2019. The study sample

originally included all courses in the College of Education with at least one registered

student during the reporting period.   For each course section included in the report, I

extracted a corresponding student data record representing each student registered in that

section as of the University census date – a fixed date for each term following the

add/drop open registration period. Census date is officially designated for each term by

the University as the moment in time used to "snap-shot" institutional data and generate

required regulatory reporting regarding registrations or enrollment.   Reported course data

included descriptive data such as: term, course department, course level, instructional

method, meeting type, course name, course number, and section number. Student registration data were reported as a single record for each student enrolled in a course section. Student registration data included student demographic data: sex (coded as Gender in the University ERP system and defined as "female," "male," or "not known"), residency (defined as "Resident" or "Non-Resident"), international reported race (calculated by the University), student level ("graduate" or "undergraduate"), degree, major, and primary college (as defined by degree and major). Student registration data also included course section data: withdrawal indicator, final grade, hours registered, and hours passed.

Student Data

International reported race is calculated for each student by the University based on answers provided by the student to three demographic questions asked at the time of first enrollment or application for admission. Each student must identify a current citizenship status of either "U.S. Citizen," "Permanent Resident," or "Non-resident Alien." Before registration is confirmed, this selection is further verified by the University using Federally and State mandated documentation of proof-of-citizenship. Optionally, the student may elect to indicate one or more races from the following list:

- American Indian or Alaskan Native
- Asian
- Black or African American
- Native Hawaiian or Other Pacific Islander
- White

Additionally, the student may indicate optional ethnic information by selecting or leaving not selected an ethnicity of "Hispanic/Latino (including Spain)".

Combined, the student's answers to these questions of citizenship, race, and ethnicity are then translated into a single indicator of race/ethnicity reported by the University in Banner as "International Reported Race." Table 3.3 details logical rules applied to the student's responses to the citizenship, race, and ethnicity questions to convert these questions into this single indicator.

Table 3.3	
Rules for Assignment of a Student's International Reported Race in Banner	
Calculated International Reported Race	Criteria for Assignment
International	Citizenship status equals "Nonresident Alien"
Hispanic/Latino (including Spain)	not International and selected ONLY Hispanic or ONLY Hispanic AND White
American Indian or Alaska Native	not International and selected ONLY this race
Asian	not International and selected ONLY this race
Black or African American	not International and selected ONLY this race
Native Hawaiian or Other Pacific Islander	not International and selected ONLY this race
White	not International and selected ONLY this race
Multiple	not International and selected two or more from above
Unknown	not International and none of the above selected

Since student grade data are reported as a letter grade in the University data, I created a calculated Numeric Letter Grade using the scale provided in Table 3.4. Grades outside of the standard letter grades (A-F) were coded with numbers greater than four, so that they could be eliminated from subsequent analyses. Final course grade was also used to calculate a Completed/Not Completed indicator. Courses in which the student received a terminal grade, including any standard letter grade, or any indicator of "P"

Table 3.4

Assigned Numeric Grade and Course Completion Indicator

Grade	Numeric Grade	Course Completed (1) / Not Completed (0)
A+	4.0	1
A	4.0	1
A-	3.7	1
B+	3.3	1
B	3.0	1
B-	2.7	1
C+	2.3	1
C	2.0	1
C-	1.7	1
D+	1.3	1
D	1.0	1
D-	0.7	1
F	0.0	1
W	10.0	0
WE	11.0	0
P	12.0	1
I	13.0	0
MG	14.0	0
R	15.0	1
AU	16.0	0

(passed) or "R" (registered) were considered as Completed (1). Any course where the student received a "W" (withdrawal), "WE" (excused withdrawal), "I" (incomplete), or "MG" (missing grade) were considered as Not Completed (0). It should be noted that since the data reported were derived from the census date, which is, by definition, after the general period when students are allowed to add or drop a registration for any reason and not incur financial or other penalties, the sample included an expected low number of withdrawals (<1%).

Course Data

Course schedule information for each Spring and Fall term, beginning Spring 2017 through Fall 2019 was extracted from Banner. This resulted in a list of 2,596 course sections offered over six terms. This data included course Instructional Method, course level, course campus, course department, course number, section number, and the total number of registered students for each section. Included in these data were courses offered by the College at the main, or primary, campus in the Northeastern United States, as well as two international, satellite campuses located in Japan and Jamaica. Because the oversight and assigned instructional method of courses offered on satellite campuses is managed entirely by the separate administration and instructional staff at those locations and these courses are enrolled entirely by students located at those locations, these courses (207 sections; 8% of the total) were considered potentially too dissimilar from the main campus courses and were, subsequently, removed from further consideration. Similarly, courses enrolling only one student (i.e., independent studies or directed research), and all thesis or dissertation writing courses (identified with course numbers 9996, 9998, and 9999) were considered substantively different in student learning outcomes and instructional design as compared to the other, more typical skills

and knowledge development courses. These courses (205 sections; 7.8% of the total) were, therefore, also eliminated from the sample.

As a College of Education includes several degree programs and courses of study designed for the professional preparation of teachers, school administrations, counselors, and other community-based support professionals; a number of courses in the College include experiential learning as either a primary or substantive portion of the course. These courses vary in range of mostly scheduled class meetings (either face-to-face, hybrid, or online) with some component of field work to primarily field work with some scheduled meeting times. These courses are identified in the University data with the Instructional Method of: "Community-based Learning," or "Non-Classroom (Field & I.S.)" and a corresponding Meeting Type of "Class," "Practicum," or "Non class meeting" It was unclear from the data provided which of these courses were organized similarly to a lecture course and which involved primarily regular field work supported by one-on-one meetings between a field supervisor and the student (essentially an independent study model). As with independent studies and research, I considered field-based pedagogies substantially different from traditional classroom-based instruction (whether delivered in-person or online), and, in order to conduct a comparison of similar groups, I eliminated these experiential learning courses (294 sections; 11.3% of the total) from the sample as well.

The calculated Numeric Letter Grade and Course Completion Indicator for each student registered in a course section were then used to calculate a Mean Letter Grade and Total Number of Completers for each of the remaining sample of 1,890 sections and these data were appended to the course section data. Additionally, the University provides an incoming cumulative Graduate GPA and Undergraduate GPA calculated for each student, depending on the student's assigned "level" (i.e., graduate or undergraduate program of study). The associated mean cumulative GPA for each course section – based

on the assigned graduate or undergraduate course level -- was calculated using student registration data and appended to the course section data.

Student Feedback Data by Course

Course feedback data for each term in the sample were extracted from the semester reports generated from the University's electronic student feedback forms (e-SFF) system and provided to the College of Education by the University's Office of Institutional Research and Assessment. These reports include a summary of the student feedback for each course section. Included in these reports, at the course level, is the total number of feedback forms generated and the total number of feedback forms completed and submitted by students. For each question, the report includes the total number of responses submitted, the number and percentage of responses submitted for each item on the five-point scale (or "Not Answered") and the calculated mean of that scale for each question at the course section, department, college, and university levels. For the purposes of this comparison study, only the section-level data were extracted from these reports.

Each course in the SFF data was coded as one of three general types: standard, recitation, and multi-instructor. Multi-instructor courses have one or more instructors designated as "primary" for the course. As only single-instructor courses were included in the extract from Banner, these SFF data were removed and disregarded. Recitations are primarily small group meetings of subsets of students simultaneously enrolled in a larger section of a lecture course. Recitations are frequently led by graduate teaching associates and other adjunct or junior faculty. I determined that recitation sections received a different SFF form with different questions than the standard course sections. In order to conduct a comparison of student satisfaction between sections, I, therefore, removed the four sections of recitation data that were included in the e-SFF report.

As defined by Hole (2006), Chi-square (χ^2) is a measure of the discrepancy between a set of observed and expected frequencies. Chi-square was used to determine if there was a significant relationship between the variables. A Cramer's V post-test is reported to indicate the effect size and importance.

Construction of the Sample

Course section data, as described above, were then merged with the remaining reported data from the University's e-SFF system. In addition to the four sections of recitation already removed, I identified 275 course sections (10.6% of the total) that appeared in the course section data but did not have any corresponding record in student feedback data[iii]. These sections were removed from the sample. An additional 31 courses were identified in the SFF data that were not listed in the course section data. I assumed these records were the result of data coding errors in the SFF system and I, therefore, excluded them from further consideration. This resulted in a sample that included a total of 1,611 course sections or 62% of the total number of course sections included in the original data extracted from Banner.

The major purpose of the research was to compare face-to-face with hybrid and online courses with specific reference to student academic success (grades), course completion rates, and student satisfaction (SFF data). As discussed, and presented in Table 3.1, the University codes the method of instruction in a more nuanced way using

---

[iii] I am not able to determine why these courses did not have corresponding student feedback data. Some known possibilities include: the sections were under-enrolled; the instructor failed to deploy or request feedback forms; or there was an error in the feedback collection process; or no students responded to the invitation to complete feedback for these courses.

several different classifications. Table 3.5 presents the Instructional Methods represented

by the 1,611 course sections in the sample as constructed.

Table 3.5	
Course Sections in the Sample by Instructional Method	
Instructional Method	Sections
Classroom	1,355
Online with some face-to-face	117
Online with required synchronous meetings	12
Online	127
Total Sections	1,611

I considered that the difference between "Online with required virtual meetings"

and "Online with some meetings required" appeared to be a more nuanced distinction

than was relevant for the purposes of this study. As a result, I combined these two groups

of Instructional Methods into a single comparison group identified as "Hybrid". This

study, then, compared three groups as defined by three levels of online learning,

including:

1. Face-to-face: including the Instructional Method of "Classroom;"
2. Hybrid: including "Online with required virtual meetings" and "Online with some meetings required"; and
3. Online: including the Instructional Method of "online."

As shown in Table 3.6, in the resulting categorized sample of the 1,611 sections

almost 84% of the course sections were taught as face-to-face ($n = 1,355$). Hybrid ($n = 129$) and online ($n = 127$) course sections each made up about 8% of the total sections

considered. I noted a large differential between the groups within the sample size: face-

to-face ($n = 1,355$) and hybrid and online ($n = 256$; or 16% of the total sections offered).

As the College historically favors traditional classroom instruction, I expected that the

number of hybrid and online learning course sections offered would be much smaller

Table 3.6

Face-to-Face, Hybrid, and Online Sections

Instructional Method	Courses	Sections	Percent of Sample
Face-to-Face	237	1355	84.11%
Hybrid	54	129	8.01%
Online	45	127	7.88%
Total Sections	336	1611	100.00%

than the number of course sections identified as "Classroom" Instructional Method. To address this issue, additional groupings and analyses were included in the research design.

Research Design

Quasi-experimental design uses non-equivalent groups that exist naturally or are already formed and is appropriate for studies where individuals cannot be randomly assigned (Trochim, 2006). The quasi-experimental design of this study involves three comparison groups as defined by Instructional Method – face-to-face, hybrid, and online teaching – and three effects:

1. student success as measured by each section's calculated Mean Letter Grade;

2. course completion as measured by the percentage calculated using the derived Total Number of Completers divided by the reported total students registered in each section; and

3. student satisfaction as measured by the means of the student responses on a five-point Likert scale, as reported by the University at the course section level for each of 11 questions on the existing University instrument used for measuring student feedback.

Additional analysis of student characteristics including sex, race/ethnicity, and residency were conducted to determine differences in student population between two broader groups of Instructional Method: classroom; and hybrid combined with online. Descriptive statistics, including means, standard deviations, and frequency distribution tables were generated to describe and summarize data in order to identify patterns.

Comparison of Student Success

Research Question #1: Is there a significant difference in student academic success, as measured by course grades, between Education courses offered through different instructional methods (online/hybrid/face-to-face)?

Data: Reported student grades for College of Education courses offered from Spring 2017 through Fall 2019 were analyzed. Means of letter grades were calculated at the course section and compared using an ANOVA for the course sections offered in different formats at the graduate and undergraduate course level.

Research Question #1a: If a difference in grades is found, does this difference remain if a student's incoming student GPA is held constant?

Data: A Mean Course GPA for each section was calculated using the associated graduate or undergraduate, as determined by the course level, GPA of each registered student. Using the resulting Calculated Course GPA for each section as a covariate, means of letter grades were compared using an ANOVA for the course sections offered in different formats at the graduate and undergraduate course level.

Comparison of Course Completion

Research Question #2: Is there a significant difference in measures of course completion between Education courses offered through different instructional methods (face-to-face/hybrid/online)?

Data: Calculated for each section: Total Number of Completers (determined by number of students receiving a terminal grade) and Percentage of Completers (Total Number of Completers as calculated divided by total number of Registered Students as reported).   An ANOVA was used to compare the means of the percent of course completers for the course sections offered in face-to-face, hybrid, and online Instructional Method.

Comparison of Student Satisfaction

Research Question #3: Is there a significant difference in reported student satisfaction with either the course materials or instruction for Education courses offered through different instructional methods (online/hybrid/face-to-face) as indicated by a course level analysis of SFFs?

Data: Each of 11 SFF questions are scored on a 5-scale Likert measure where 1 = *strongly disagree*, 2 = *disagree*, 3 = *neutral*, 4 = *agree*, and 5 = *strongly agree*. Means for each question are reported at the course section   A repeated measures ANOVA was computed comparing the three groups of course sections categorized by Instructional Method. As the measurements (question answers) are repeated such that random factors that cause a particular value to be high or low are not addressed and the order of the questions are not randomized, the assumption of sphericity is violated (Bathke, Schabenberge, Tobias, & Madden, 2009).   The Greenhouse-Geisser correction was therefore applied.

Consideration of Student Characteristics

Research Question #4: Are there significant differences in student characteristics of sex, race/ethnicity, and residency between students who enroll in face-to-face as compared to online or hybrid courses?

Data: Because of the disparity between the group sizes that included a large percentage of face-to-face course sections in the sample, to answer this question I

combined online and hybrid courses into a single group and designated that as "online." Characteristics of students registered for each course section were identified in the student registration data including: student sex: reported as "female", "male", or "not known;" the University calculated student race/ethnicity reported in International Reported Race; and the University determined student residency status reported as "resident" or "non-resident". As defined by Hole (2006), Chi-square (χ^2) is a measure of the discrepancy between a set of observed and expected frequencies. Chi-square was used to determine if there was a significant relationship between the variables. A Cramer's V post-test is reported to indicate the effect size and importance.

Limitations of the Design

This study relies on existing data to consider relationships between course instructional method and measures of student success, completion, and satisfaction. Students self-selected and were not randomly assigned into different course delivery methods. Factors influencing a student's decision to enroll in an online course versus an on-site course, including considerations such as time of day the course is offered, frequency of course offerings, number of sections offered in different formats, or whether the course is required in the student's program will not be considered in this study and may have influenced the level of student success and satisfaction. In addition, except for the limited consideration of the completion rate of each student, additional influencing factors on student success and satisfaction, such as the student's prior academic history and comfort-level with technology are not considered in this analysis. Influencing factors that may be related to the characteristics of the course instructor, such as the age, sex, and seniority of the instructor, are also not included in these data for consideration. Equally, while some decisions have been made to remove from consideration courses that are

fundamentally different in format and delivery, differences between course design of different sections of the same course, including content and assessments of student learning, are not known and cannot be considered. Since several factors and independent variables not considered in this study are open to further speculation as to their potential level of influence, this study does not seek to discuss why differences may or may not be present between courses offered in different instructional methods. Further research will be required to address causation.

Role of the Researcher and Ethical Considerations

I am currently an administrator and data steward in the College of Education that this study considers. I also provide a role in technology support and adoption by instructors in the College. In my current role, I was able to gather a majority of the data needed for analysis using existing access to reporting tools, such as IBM Cognos. I de-identified and otherwise modified the data described in this methodology to remove personally identifiable data from all extracted reports.

This study is secondary research derived from data previously collected by the University for the purposes of educational research and programmatic improvement and, as such, represents a minimal risk to subjects included for consideration. SFF data, as reported by the institution, are aggregated by the course section and do not include individual student responses. Ethical considerations to this study are limited to any harm that may come of providing commentary on analysis of existing data and from the disclosure of academic records. All personally identifiable information was removed from the data prior to analysis, reducing the possibility of identifying any particular student or instructor.

Summary

This chapter outlined the methodological approach of the this quasi-experimental design study. The study compared measures of student success, course completion rates, and student satisfaction between different instructional methods for graduate and undergraduate courses offered in a College of Education at a large, urban, Northeastern university during the Spring 2017 through the Fall 2019 semesters. Data were obtained through existing University databases of student records including course schedules, student academic history, and student feedback forms. Data were analyzed using ANOVA and multiple tests to determine levels of difference for course delivery methods using student attributes including sex, race/ethnicity, and residency status.

CHAPTER 4: RESULTS

The purpose of the current study was to review the data connected with students who take education studies courses in different delivery formats. Specifically, the data focused on both course performance and what, if any, course delivery formats correlate with performance and student satisfaction as evidenced by final grade and student evaluations of the course. In this chapter, characteristics of the data, the courses, and the sample are discussed as well as the findings of the data analyses.

Description of the Sample

In Chapter 3, I detailed the steps completed to derive the sample for this study. To assist in providing a comprehensive description of the sample, a brief review of the major components of that process, along with details regarding the number of records and other relevant characteristics of the sample as of each step is included here.   As explained in Chapter 3, data for this study were obtained from the following sources:

- Banner: course schedule information including course instructional method and meeting type, course level, and total number of registered students.
- Banner student course registration data including student registration status, final grades and cumulative GPA, demographics (sex, race/ethnicity/residency status), major, and primary college of study.
- Office of Institutional Research and Assessment e-SFF system reports including data from student feedback forms (SFF) reported at the course-level for each of 11 standard questions.

The potential sample derived from the student records and registration data originally included 2,596 course sections taught in the College of Education from Spring

2017 through Fall 2019. To make the analyses adequately represent typical instruction in the College, I made several decisions to construct the final course data set:

1. all independent study courses were eliminated;
2. all courses associated with thesis or dissertation writing (i.e., 9996, 9998 and 9999), including research apprenticeship, were eliminated since these courses are essentially independent studies;
3. all courses offered in Japan were eliminated; and
4. all courses offered in Jamaica were eliminated.

I next evaluated courses identified as field-study, practicum, or internship. These courses vary in range of mostly scheduled class meetings (either face-to-face or online) with some component of field work to primarily field work with some scheduled meeting times. It was unclear from the data provided which of these courses were organized similarly to a lecture course and which involved primarily regular field work supported by one-on-one meetings between the field supervisor and the student (essentially an independent study model). As a result, these courses were also eliminated from the sample. Table 4.1 presents a summary of the distribution of the 2,596 course sections, the exclusions applied, and the resulting potential sample of 1,890 course sections.

After the initial construction of the sample, as described above, the course data included 295 different courses taught in the College with a total of 1,890 unique sections. I then merged these course section data with SFF data available for each term. I determined that "Recitation" sections received a different SFF form with different questions than the other courses. In order to conduct a comparison of student satisfaction between sections, I, therefore, excluded the four sections of recitation that received different SFF forms. I identified an additional 279 course sections that appeared in the

Table 4.1

Courses in the Sample and Exclusions

| Term | All Courses | Excluded Course Types | | | | | Total Included |
		Dissertation, Thesis, and Research	Practicum, Field, and Internship	Independent Study	Jamaica	Japan	
2017 Spring	412	29	43	5	5	33	297
2017 Fall	449	29	46	3	5	32	334
2018 Spring	425	25	51	4	5	34	306
2018 Fall	434	29	48	5	4	27	321
2019 Spring	423	28	54	6	3	30	302
2019 Fall	453	32	52	10	2	27	330
Total	2596	172	294	33	24	183	1890

course data but did not have any corresponding student feedback data[iv]. For parity of analyses, these sections were also removed from further consideration. As shown in Table 4.2, the resulting sample included 277 different courses with a total of 1,611 unique sections.

Table 4.2			
Courses in the Sample with Recitation and Missing SFF Excluded			
Term	Courses	Sections	Percent of Sample
2017 Spring	148	259	16.08%
2017 Fall	157	289	17.94%
2018 Spring	155	263	16.33%
2018 Fall	157	272	16.88%
2019 Spring	153	257	15.95%
2019 Fall	156	271	16.82%
All Terms	277[v]	1611	100.00%

An additional 31 courses were identified in the SFF data that were not listed in the records and registration data. I assumed these records were the result of data coding errors in the SFF data and I, therefore, excluded them from further consideration.

[iv] I am not able to determine why these courses did not have corresponding student feedback data. Some known possibilities include: the sections were under-enrolled; the instructor failed to deploy or request feedback forms where manual intervention was required; there was an error in the feedback collection process; or no students responded to the invitation to complete feedback for these courses.

[v] This number is not a sum of the column, but, rather, represents the total number of unique courses taught over the entire reporting period.

Classifying Method of Instruction

The major purpose of the research was to compare face-to-face with hybrid and online courses with specific reference to student academic success (grades), course completion rates, and student satisfaction (SFF data). The University codes the method of instruction in a more nuanced way using several different classifications. This is presented in Table 4.3 using the original course data set identified before being merged with SFF data.

Table 4.3			
Sections by Instructional Method			
Instructional Method	Courses	Sections	Percent of Sample
Classroom	251	1585	84.04%
Community Based Learning	4	7	0.37%
Non-Classroom (Field & I.S.)	16	27	1.43%
Online with required virtual meetings	8	12	0.64%
Online with some meetings required	51	125	6.63%
Online	46	130	6.89%
Total Sections	376	1886	100.00%

"Community Based Learning", and "Non-Classroom (Field & I.S.)" courses include experiential, site-based learning that introduces a number of unknown variables, including: site type, location, and percentage of learning occurring on site. As with the other practicum, field, and internship courses previously eliminated from the course data sample, I considered these variables outside the scope of like-course comparison for this study and, therefore, decided to eliminate these additional 34 sections from the data.

As discussed in Chapter 3, Instructional Method can be used to assign a level of online learning representing the extent of online learning included in a given course as continuum from "entirely face-to-face" to "entirely online" with a number of variations in between. As shown in Table 4.4, as the number increases the extent to which the course is online increases. After the sample was constructed as described above, I determined that there were only four levels of online learning represented in the remaining data set. The distribution is shown in Table 4.5.

Table 4.4	
Level of Online Learning	
Instructional Method	Level of "Online"
Online	13
Online with required synchronous meetings	12
Online with some face-to-face	11
Video conference	10
Classroom	9
Classroom & Applied Experience	8
Recitation	7
Studio	6
Laboratory	5
Learning Community	4
Community Based Learning	3
Non-Classroom Field Experience	2
Supervised Experience	1

Given the small number of sections identified as "Online with required virtual meetings," the difference between "Online with required virtual meetings" and "Online with some meetings required" appeared to be a more nuanced distinction than was relevant for the purposes of this study.   I decided to examine three levels of online learning, including: "Classroom" defined as face-to-face; "Online with required virtual

Table 4.5

Instructional Methods in Sample

Instructional Method	Sections
Classroom	1,355
Online with some meetings required	117
Online with required virtual meetings	12
Online	127
Total Sections	1,611

meetings" and "Online with some meetings required" combined and identified as

"Hybrid"; and "Online" as the designation for online learning. As shown in Table 4.6, in

the resulting categorized sample of 1,611 sections almost 84% of courses were taught as

face-to-face. Hybrid and online courses each made up about 8% of the total course

sections considered.

Table 4.6

Face-to-Face, Hybrid, and Online Sections

Instructional Method	Courses	Sections	Percent of Sample
Face-to-Face	237	1355	84.11%
Hybrid	54	129	8.01%
Online	45	127	7.88%
Total Sections	336	1611	100.00%

As I noted in Chapter 3, there is a large differential in the groups within the

sample size: face-to-face ($n = 1,355$) and hybrid and online ($n = 256$) or 16% of the total

sections in the sample. The College historically favors traditional classroom instruction

resulting in fewer online learning course offerings in the sample. Therefore, additional

groupings and analyses were conducted, as detailed below, in consideration of the

disparity between the size of the comparison groups in sample.

Student Registration Data

For each course section included in the sample, I extracted a student data record representing each student registered in that section as of the University census date – a fixed date for each term following the add/drop open registration period. Census date is officially designated for each term by the University as the moment in time used to "snap-shot" data and generate required regulatory reporting regarding registrations or enrollment. Using the census date of each term, for the 1,611 course sections included in the sample, a total of 33,948 student registration data points were identified. Each registration record represents a single student enrollment in one of the College's course sections. As shown in Table 4.7, the total student registrations records included 25% graduate students and 75% undergraduate students.

Table 4.7

Student Registration Records in the Sample

Student Registrations	Graduate		Undergraduate		Total
Term	Registrations	Percent	Registrations	Percent	Registrations
2017 Spring	1,408	26%	4,047	74%	5,455.00
2017 Fall	1,588	25%	4,721	75%	6,309.00
2018 Spring	1,329	25%	3,969	75%	5,298.00
2018 Fall	1,481	25%	4,394	75%	5,875.00
2019 Spring	1,273	24%	3,947	76%	5,220.00
2019 Fall	1,504	26%	4,287	74%	5,791.00
Total Registrations	8,583	25%	25,365	75%	33,948.00

All courses were offered in the College of Education; however, students from across the University are permitted to register in most courses. Table 4.8 shows the

Table 4.8

Student Registrations by Student Major College

Student Majors	Face-to-Face		Hybrid		Online		Total	
Major College	Registered	Percent	Registered	Percent	Registered	Percent	All Instructional Types	Percent
Education	19,245	65.36%	1,883	95.54%	1,959	77.37%	23,087	68.01%
Liberal Arts	1,860	6.32%	13	0.66%	113	4.46%	1,986	5.85%
Business & Management	1,866	6.34%	2	0.10%	44	1.74%	1,912	5.63%
Science & Technology	1,418	4.82%	1	0.05%	56	2.21%	1,475	4.34%
Public Health	1,221	4.15%	36	1.83%	60	2.37%	1,317	3.88%
University Studies	995	3.38%		0.00%	141	5.57%	1,136	3.35%
Media & Communication	930	3.16%	2	0.10%	39	1.54%	971	2.86%
Art& Architecture	530	1.80%	25	1.27%	71	2.80%	626	1.84%
Engineering	370	1.26%		0.00%	8	0.32%	378	1.11%
Sport, Tourism & Hospitality Management	334	1.13%	5	0.25%	6	0.24%	345	1.02%
Theater, Film & Media Arts	322	1.09%	1	0.05%	7	0.28%	330	0.97%
Music & Dance	218	0.74%	1	0.05%	17	0.67%	236	0.70%
Social Work	136	0.46%	1	0.05%	9	0.36%	146	0.43%
Medicine		0.00%	1	0.05%	2	0.08%	3	0.01%
Total Registrations	29,445	100%	1,971	100%	2,532	100%	33,948	100%

distribution of student registrations by major with almost 68% of registrations by students with a major program in the College of Education and no more than 6% of student majors from any one of the other thirteen schools and colleges in the University. Given the lack of diversity of student registrations from different colleges within the sample, I decided not to conduct any further analyses by student major or college.

Tables 4.9 – 4.12 summarize characteristics of the students in the sample. The sample included 75% undergraduate students and 25% graduate students. The majority of students are white (60%), female (69.7%), and classified by the University as in-state residents (78%). For graduate students, 15.7% were in the first or second year of their program, having completed less than 17 credit hours. Undergraduate students were distributed more evenly across the Freshmen, Sophomore, Junior, and Senior classes with a smaller number of fifth-year Seniors.

Table 4.9

Student Registrations by Sex

Sex	Students	Percent of Sample
Female	23,666	69.71%
Male	10,145	29.88%
Not Disclosed	137	0.40%
Total Registrations	33,948	100.00%

Table 4.10

Student Registrations by Residency Status

Residency of Students	Students	Percent of Sample
Pennsylvania Resident	26,596	78.34%
Non-Pennsylvania Resident	7,341	21.62%
Undeclared	11	0.03%
Total Registrations	33,948	100.00%

Table 4.11

Student Registrations by Reported Race

Reported Race	Students	Percent of Sample
White	20,561	60.57%
African American	5,345	15.74%
Asian	2,341	6.90%
Hispanic	2,254	6.64%
Unknown	1,152	3.39%
International	1,152	3.39%
Multi-Racial	1,061	3.13%
Other	82	0.24%
Total Registrations	33,948	100.00%

Table 4.12

Student Registrations by Level

Level of Students	Students	Percent of Sample
Graduate	8,583	25.28%
Graduate 0 to 17 Credits	5,340	62.22%
Graduate 18 to 35 Credits	1,944	22.65%
Graduate 36 to 53 Credits	918	10.70%
Graduate 54+ Credits	381	4.44%
Undergraduate	25,365	74.72%
Freshman 0 to 29 Credits	5,284	20.83%
Sophomore 30 to 59 Credits	4,801	18.93%
Junior 60 to 89 Credits	6,983	27.53%
Senior 90 to 119 Credits	6,548	25.82%
Senior/Fifth Year 120+ Credits	1,745	6.88%
Undergraduate Non-Degree	4	0.02%
Total Registrations	33,948	100.00%

Approach to Analyses and Benchmarks

The data to answer the research questions came primarily from course data combined with the student feedback form data. To analyze these data, several different tests were used to produce the results as presented for each research question. The two primary techniques included the analysis of variance (ANOVA) with repeated measures and chi square. These analyses were performed to analyze the relationship between measures of student success, course completion, and student satisfaction. Differences that were significant at the .05 level are discussed. Because of the strength of the sample size, it is important to avoid reporting results that are statistically significant but realistically meaningless (Tabachnick & Fidell, 2000).

For the ANOVAs presented, partial eta squared is provided to indicate the effect size or the ratio of variance associated with or accounted for by the main effects or interaction. The benchmarks for Partial eta squared, based on conventions for defining effect sizes in social science research and adapted from Murphy and Myors (2004) are presented in Table 4.13.

Table 4.13

Benchmarks for Partial Eta Squared

Rating	Description
$0.01 \leq \geq 0.05$	Small
$0.051 \leq \geq 0.13$	Medium
≥ 0.14	Large

The Cramer's V test was performed to analyze the strength of the relationship between characteristics of students and instructional method assessed though chi square. The benchmarks for Cramer's V analysis (Zaiontz, 2014) are presented in Table 4.14.

Table 4.14

Benchmark for Cramer's V

Rating	Description
0.07 ≤ ≥ 0.2	Weak relationship
0.21 ≤ ≥ 0.34	Moderate relationship
≥ 0.35	Strong relationship

Analyses of the Data for the Research Questions

This study addresses several key questions that relate to students' achievement, persistence, and satisfaction when comparing face-to-face, hybrid, and online learning environments. Students' performance, as measured by the grade earned in the class, was used to determine whether students perform better or worse in education studies courses offered using different instructional methods. I included an analysis of the students' cumulative GPA as a measure of student preparedness that may act as a secondary factor to instructional method that impacts student success. Second, this study considers differences in student persistence, as measured by the presence or absence of codes indicating student withdrawal, failure to complete the course requirements, or unexplained stop-out before the end of the term. Third, through a review of student feedback collected for each course section, this study explores any difference in students' perceptions of their course, instructor, or learning as a function of the method of instruction. Finally, as the students self-select the delivery method for any course, it was important to determine if there were any defining characteristics that may influence a student choosing (or not choosing) a hybrid or online course, when available, over courses offered using traditional classroom instruction.

Comparison of Student Success

> *Research Question #1*: Is there a significant difference in student academic success, as measured by course grades, between Education courses offered through different instructional methods (online/hybrid/face-to-face)?

This question focuses on student achievement outcomes when comparing face-to-face, hybrid, and online courses. To address this question, an ANOVA was conducted to assess whether students performed better in one or the other course delivery methods, as measured by final course grade. In consideration of the assumed difference in the learning outcomes and characteristics of the students, the sample was split not only by instructional method, but also by associated course level: graduate ($n = 569$) and undergraduate ($n = 1042$). Table 4.15 presents the means and the standard deviations by the numeric translation of the letter grade received.

Table 4.15

Means and Standard Deviation of Course Grades

Instructional Methods by Level		Mean	Std. Deviation	N
Face-to-Face	Graduate	3.800	0.223	403
	Undergraduate	3.578	0.342	952
Hybrid	Graduate	3.794	0.210	105
	Undergraduate	3.346	0.394	24
Online	Graduate	3.782	0.229	61
	Undergraduate	3.386	0.374	66
Overall	Graduate	3.797	0.221	569
	Undergraduate	3.560	0.350	1,042
	Total	3.644	0.330	1,611

I note, here, that there are large differences in sample size among the groups. This is unavoidable because of the relatively small number of hybrid and online courses. While none of the sample sizes are "small" in an absolute sense, the issue of different sample sizes affects almost all of the analyses. Since the power of an ANOVA is based

on the smallest sample, the effect is typically to make it harder to find statistical significance. In addition, post-hoc tests are not well designed to handle large sample size differences. As a consequence, where there are significant differences, post hoc tests will only be reported in one case where the sample size issue was not relevant.

As shown in Table 4.16, both main effects and the interaction are statistically significant with small effect sizes. Of most interest for this study, it is evident from Figure 4.1 that students in the face-to-face courses obtain higher grades than the students in either the hybrid or online courses. Although, as already mentioned, the effect size is very small (.014). The difference is greater for undergraduates than for graduate students with undergraduates lower for hybrid than online and both instructional methods lower than face-to-face. For Research Question 1, it is clear that the means for hybrid and online sections are significantly lower for undergraduates, but the difference overall, and within the course level, is too small to be considered meaningful.

Table 4.16

ANOVA Summary for Course Grades and Instructional Method

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Instructional Method	2.127	2	1.063	11.261	0.000	0.014
Course Level	13.166	1	13.166	139.431	0.000	0.080
Instructional Method * Level	1.649	2	0.824	8.730	0.000	0.011
Error	151.555	1605	0.094			
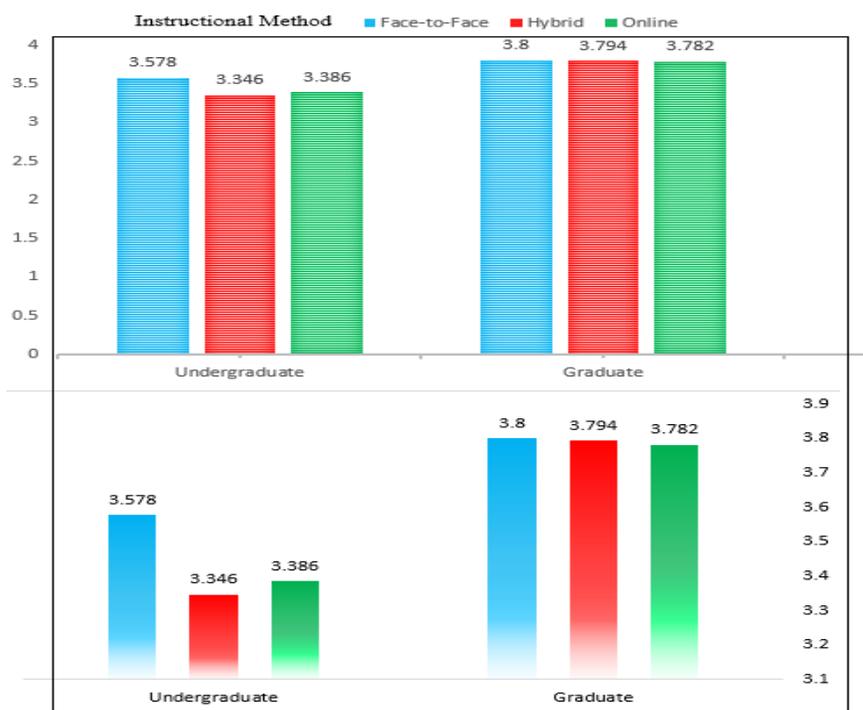
Figure 4.1 Graph of Means of Average Letter Grades

Question 1a: If a difference in grades is found, does this difference remain if a student's incoming student GPA is held constant?

As discussed above, the difference in student success may be attributed to factors other than the instructional method. For example, stronger students, with a history of academic success, may be less impacted by the delivery format. To consider the influence of student academic preparedness, the analysis presented above was run again using student cumulative GPA as the covariate. The results are presented in Figure 4.2 and Table 4.17.

The results remained essentially the same, with a slightly greater difference between face-to-face and hybrid and online for graduate students, along with a slight decrease in the effect size (.008). This implies that instructional method may have a greater impact on student success, even for stronger students. However, the effect size is so small it may not be considered meaningful.
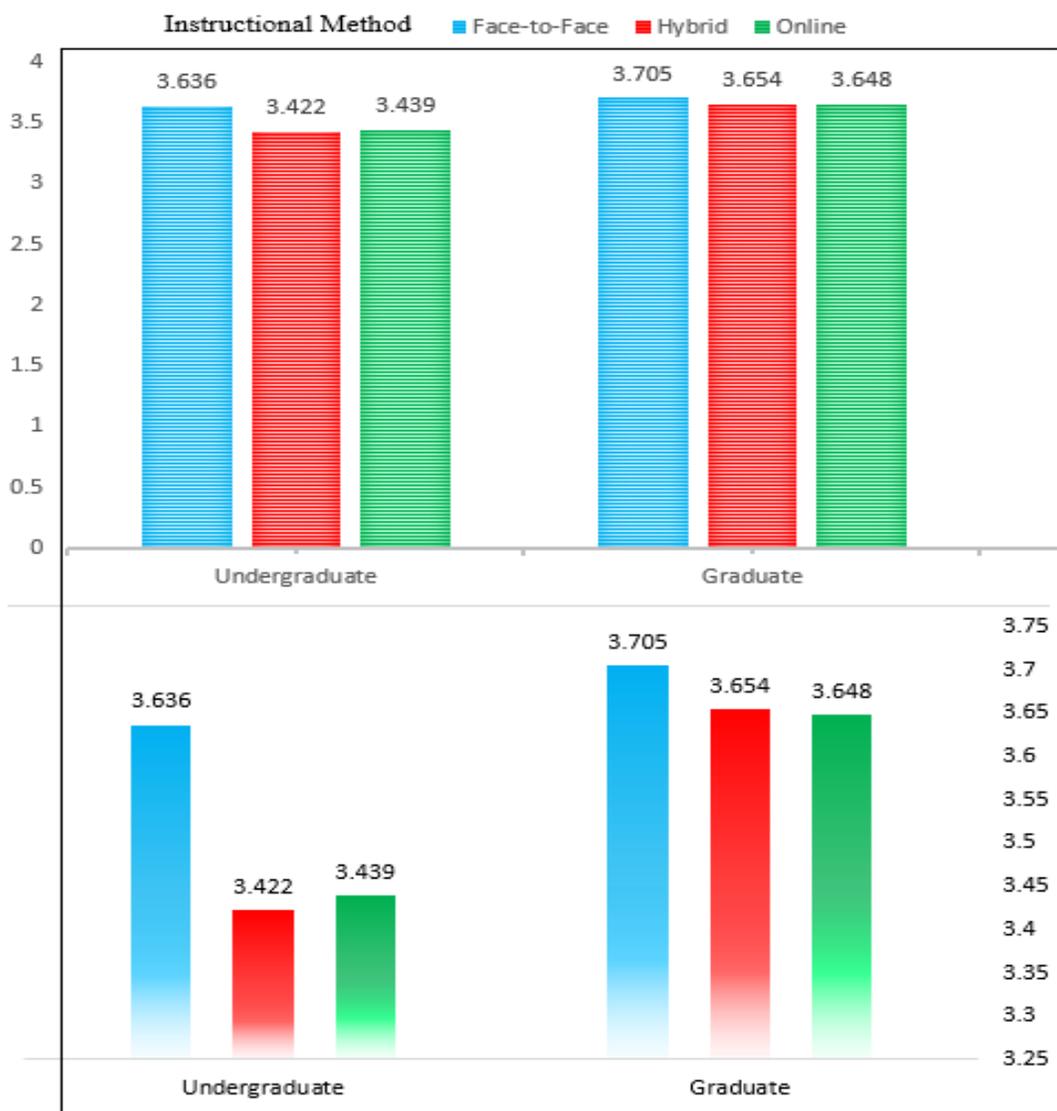
Figure 4. 2 Graph of Means of Average Letter Grades with Student GPA Covariate

Table 4.17

ANOVA Summary with Student GPA Covariate

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Student GPA	35.694	1	35.694	494.152	0.000	0.236
Instructional Method	2.903	2	1.452	20.096	0.000	0.024
Course Level	2.752	1	2.752	38.103	0.000	0.023
Instructional Method * Level	0.968	2	0.484	6.704	0.001	0.008
Error	115.861	1604	0.072			

Comparison of Course Completion

Research Question #2: Is there a significant difference in measures of course completion between Education courses offered through different instructional methods (face-to-face/hybrid/online)?

To answer this question, students who obtained a withdrawal (W), missing grade (MG) or incomplete (I) were considered to have not completed the course. I calculated a percentage of course completers as the number of completed students divided by the total number of registered students for each course. The means and standard deviation of the percent of completed by instructional method, again broken out between graduate and undergraduate courses, is presented in Table 4.18.

Analysis of the means of percent completed was then conducted and the results are presented in Table 4.19 and Figure 4.3. As with the analysis of student success above, effect size (.014). The difference for graduate courses was smaller than with undergraduate courses, with online undergraduate courses having the lowest student completion. Notably, student completion for graduate courses was almost identical across instructional methods; however, face-to-face was the lowest with higher means of completion for hybrid and online, respectively. As with the analysis of student success above, the difference in student completion by instructional methods was significant with

Table 4.18

Means and Standard Deviation of the Percent of Course Completers

Instructional Methods		Mean	Std. Deviation	N
Face-to-Face	Graduate	0.982	0.038	403
	Undergraduate	0.981	0.038	952
	Total	0.982	0.038	1355
Hybrid	Graduate	0.985	0.035	105
	Undergraduate	0.964	0.045	24
	Total	0.981	0.038	129
Online	Graduate	0.987	0.031	61
	Undergraduate	0.954	0.061	66
	Total	0.970	0.052	127
Total	Graduate	0.983	0.037	569
	Undergraduate	0.979	0.041	1042
	Total	0.981	0.039	1611

Table 4.19

ANOVA Summary with Percent of Student Completers Covariate

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Instructional Method	0.017	2	0.009	5.750	0.003	0.007
Course Level	0.034	1	0.034	22.610	0.000	0.014
Instructional Method * Level	0.035	2	0.017	11.393	0.000	0.014
Error	2.438	1605	0.002			

a small effect size (.014). The difference for graduate courses was smaller than with

undergraduate courses, with online undergraduate courses having the lowest student

completion. Notably, student completion for graduate courses was almost identical

across instructional methods; however, face-to-face was the lowest with higher means of

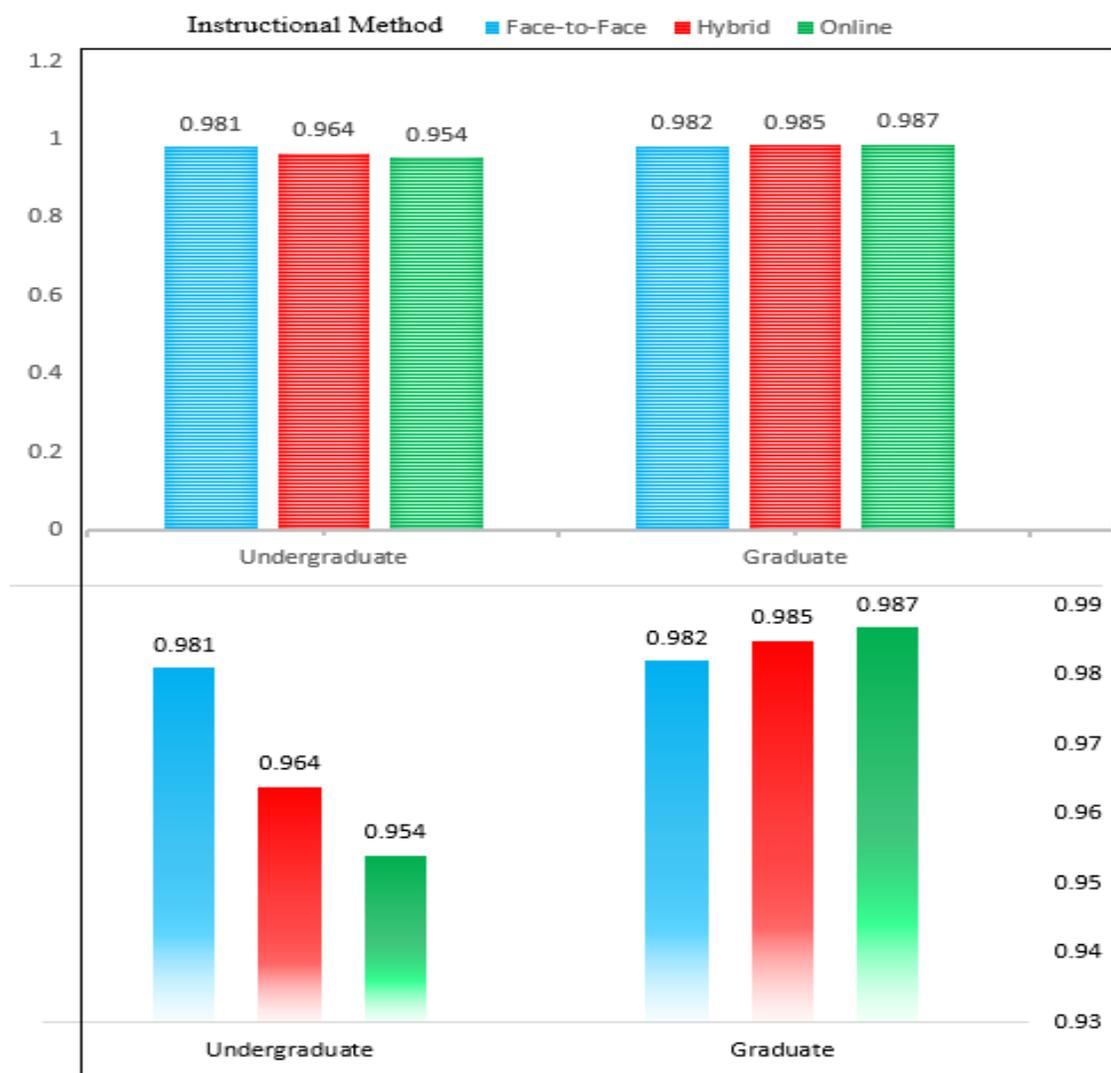completion for hybrid and online, respectively.

Figure 4.3 Graph of Means of the Percent of Student Completers

.

Comparison of Student Satisfaction

Research Question #3: Is there a significant difference in reported student satisfaction with either the course materials or instruction for Education courses offered through different instructional methods (online/hybrid/face-to-face) as indicated by a course level analysis of SFFs?

To address this question, I used the reported mean for each course section derived from the responses to each of the 11 questions on the student feedback form. As a reminder, the 11 common questions on the SFF form are presented in Table 4.20.

As I mentioned in Chapter 3, the University does not publish historical or trend completion rates for the SFF forms. At the time of this study, the University claimed that 43% of feedback forms had been completed for the semester in which this study was conducted. However, as Table 4.21 shows, the response rate for the courses in the sample was much higher overall (62%) than the available University response rate with more responses received for graduate level courses (67%) than undergraduate courses (59%). It is worth noting that the response rate for online courses was the lowest for both

Table 4.20
Questions on the Student Feedback Form
Q1: I came well prepared for class
Q2: The instructor clearly explained the educational objectives of the course
Q3: The instructor was well organized and prepared for class
Q4: The instructor was conscientious in meeting class and office hour responsibilities
Q5: The instructor promoted a classroom atmosphere in which I felt free to ask questions
Q6: The instructor provided useful feedback about exams, projects, and assignments
Q7: So far the instructor has applied grading policies fairly
Q8: The instructor taught this course well
Q9: The course content was consistent with the educational objectives of the course
Q10: The course increased my ability to analyze and critically evaluate ideas, arguments and points of view
Q11: I learned a great deal in this course

Table 4.21

Rates of Student Completion and SFF Forms Submitted

Students	Registered	Completed	Rate of Completion	SFFs Submitted	SFFs Response Rate
Graduate	8,903	8,755	98%	6,201	67%
Face-to-Face	6,462	6,349	98%	4,469	70%
Hybrid	1,470	1,449	99%	1,002	63%
Online	971	957	99%	550	59%
Undergraduate	25,045	24,530	98%	15,031	59%
Face-to-Face	22,983	22,560	98%	14,082	60%
Hybrid	501	483	96%	260	50%
Online	1,561	1,487	95%	689	44%
Total Registrations	33,948	33,285	98%	21,232	62%

graduate and undergraduate courses. Undergraduate online course completion, at 44%, was both the lowest among the groups and the only group that was similar to the overall response rate as available from the University.

An analysis of the response rates was completed with a significant difference as a function of instructional method and a significant interaction between instructional method and the questions.   However, the interaction between instructional method, course level, and the questions was not significant. These results are presented in Appendix A, but not considered relevant for further discussion here.

A repeated measures ANOVA was then computed comparing online, hybrid, and face-to-face courses on the mean SFF rating for each of the 11 questions on the form. The repeated measures ANOVAs were run separately for undergraduate courses versus

graduate courses. As the measurements (question answers) are repeated such that random factors that cause a particular value to be high or low are not addressed and the order of the questions are not randomized, the assumption of sphericity is violated (Bathke, Schabenberge, Tobias, & Madden, 2009).   The Greenhouse-Geisser correction was therefore applied. The means and standard deviations are presented in Table 4.22, along with the mean response rate for each question, which is similar across the questions. The ANOVA results are present in Table 4.23. As shown in Table 4.23, the between subjects main effects for instructional method and course level are not signficant, nor is the interaction of instructional method with course level. The within-subject effects of SFF questions with instructional method and SFF questions with course level are significant with small effects. However, there is no significant within-subject effect for the interaction of SFF questions, instructional method, and course level.

Figures 4.4 and 4.5 present the graphs of the means for the student feedback form questions at the graduate and undergraduate course level, respectively. As the graphs show, for graduate and undergraduate courses, the student feedback received is similar across instructional method for all questions with slightly higher means for online graduate courses and lower means for hybrid undergraduate courses. As there were perceivable differences on some questions, specifically three, four, six and eleven, additional post-hoc tests were done, and significant differences were found only between face-to-face and hybrid on question six. Specifically, students in hybrid courses, as compared to students in face-to-face courses, perceive that the instructor did not provide useful feedback about exams and courses. Significant differences between the means were not identified for any question when analyzed separately at the graduate and undergraduate level.

Table 4.22

Questions on SFF: Means and Standard Deviation

SFF Question	Instructional Method	Level	Mean	Std. Deviation	*N*	Mean Response Rate
Q1	Face-to-Face	Graduate	4.262	0.291	403	99%
		Undergraduate	4.359	0.280	950	99%
	Hybrid	Graduate	4.292	0.347	105	99%
		Undergraduate	4.396	0.294	24	99%
	Online	Graduate	4.379	0.310	61	95%
		Undergraduate	4.379	0.309	66	95%
Q2	Face-to-Face	Graduate	4.348	0.552	403	99%
		Undergraduate	4.422	0.514	950	100%
	Hybrid	Graduate	4.412	0.482	105	99%
		Undergraduate	4.404	0.493	24	100%
	Online	Graduate	4.464	0.383	61	98%
		Undergraduate	4.362	0.435	66	100%
Q3	Face-to-Face	Graduate	4.360	0.629	403	98%
		Undergraduate	4.447	0.567	950	99%
	Hybrid	Graduate	4.334	0.618	105	97%
		Undergraduate	4.488	0.396	24	99%
	Online	Graduate	4.454	0.507	61	97%
		Undergraduate	4.377	0.422	66	97%
Q4	Face-to-Face	Graduate	4.543	0.443	403	99%
		Undergraduate	4.521	0.457	950	98%
	Hybrid	Graduate	4.521	0.491	105	99%
		Undergraduate	4.621	0.248	24	99%
	Online	Graduate	4.528	0.418	61	88%
		Undergraduate	4.432	0.371	66	88%
Q5	Face-to-Face	Graduate	4.618	0.366	403	99%
		Undergraduate	4.575	0.422	950	98%
	Hybrid	Graduate	4.625	0.391	105	97%
		Undergraduate	4.492	0.420	24	97%
	Online	Graduate	4.567	0.392	61	90%
		Undergraduate	4.433	0.446	66	90%

SFF Question	Instructional Method	Level	Mean	Std. Deviation	N	Mean Response Rate

Table 4.22

(continued)

SFF Question	Instructional Method	Level	Mean	Std. Deviation	N	Mean Response Rate
Q6	Face-to-Face	Graduate	4.292	0.617	403	99%
		Undergraduate	4.370	0.553	950	99%
	Hybrid	Graduate	4.167	0.662	105	99%
		Undergraduate	4.263	0.401	24	99%
	Online	Graduate	4.333	0.573	61	99%
		Undergraduate	4.335	0.452	66	99%
Q7	Face-to-Face	Graduate	4.469	0.506	403	96%
		Undergraduate	4.508	0.452	950	98%
	Hybrid	Graduate	4.454	0.427	105	95%
		Undergraduate	4.379	0.521	24	99%
	Online	Graduate	4.548	0.424	61	98%
		Undergraduate	4.488	0.373	66	98%
Q8	Face-to-Face	Graduate	4.290	0.621	403	100%
		Undergraduate	4.384	0.582	950	100%
	Hybrid	Graduate	4.289	0.593	105	99%
		Undergraduate	4.292	0.579	24	99%
	Online	Graduate	4.323	0.547	61	99%
		Undergraduate	4.271	0.477	66	99%
Q9	Face-to-Face	Graduate	4.434	0.511	403	99%
		Undergraduate	4.500	0.447	950	98%
	Hybrid	Graduate	4.487	0.442	105	98%
		Undergraduate	4.475	0.342	24	98%
	Online	Graduate	4.521	0.382	61	99%
		Undergraduate	4.474	0.351	66	98%
Q10	Face-to-Face	Graduate	4.300	0.534	403	99%
		Undergraduate	4.328	0.512	950	99%
	Hybrid	Graduate	4.276	0.558	105	99%
		Undergraduate	4.283	0.491	24	99%
	Online	Graduate	4.270	0.475	61	99%
		Undergraduate	4.256	0.386	66	100%

Table 4.22						
(continued)						
SFF Question	Instructional Method	Level	Mean	Std. Deviation	*N*	Mean Response Rate
Q11	Face-to-Face	Graduate	4.272	0.581	403	100%
		Undergraduate	4.299	0.543	950	99%
	Hybrid	Graduate	4.251	0.558	105	100%
		Undergraduate	4.217	0.477	24	100%
	Online	Graduate	4.280	0.444	61	100%
		Undergraduate	4.285	0.394	66	100%

| Table 4.23				
Repeated Measures Summary of Results				
Between Subjects Effects				
Source	Mean Square	F	Sig.	Partial Eta Squared
Instructional Method	0.210	0.102	0.903	0.000
Course Level	0.023	0.011	0.915	0.000
Instructional Method * Course Level	1.602	0.779	0.459	0.001
Error Between	2.055			
Within Subjects Greenhouse-Geisser				
Source	Mean Square	F	Sig.	Partial Eta Squared
SFF Questions	8.508	67.335	0.000	0.040
SFF Questions * Instructional Method	0.423	3.351	0.000	0.004
SFF Questions * Course Level	0.462	3.659	0.002	0.002
SFF Questions * Course Level * Instructional Method	0.163	1.289	0.228	0.002
Error Within	0.126			
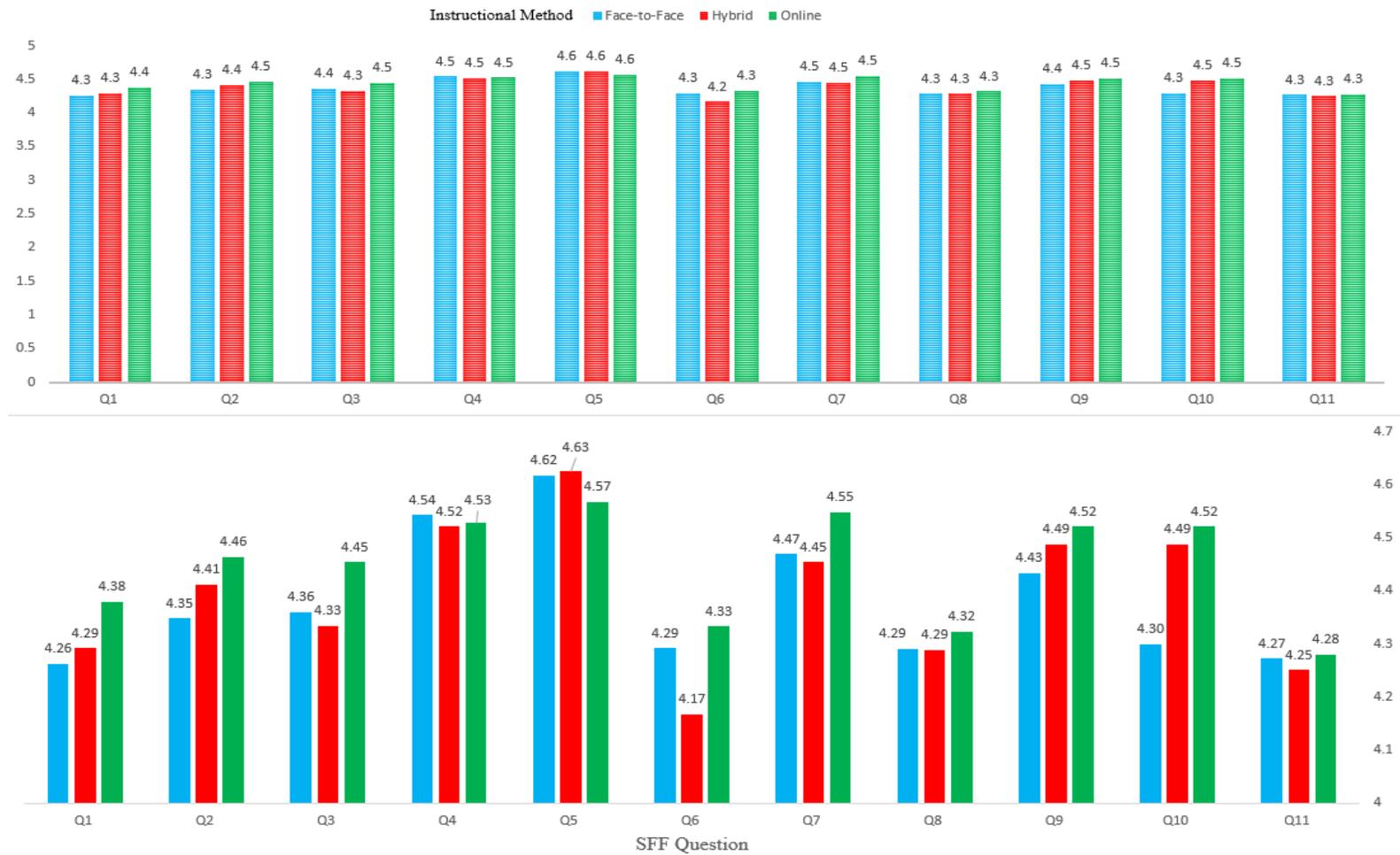
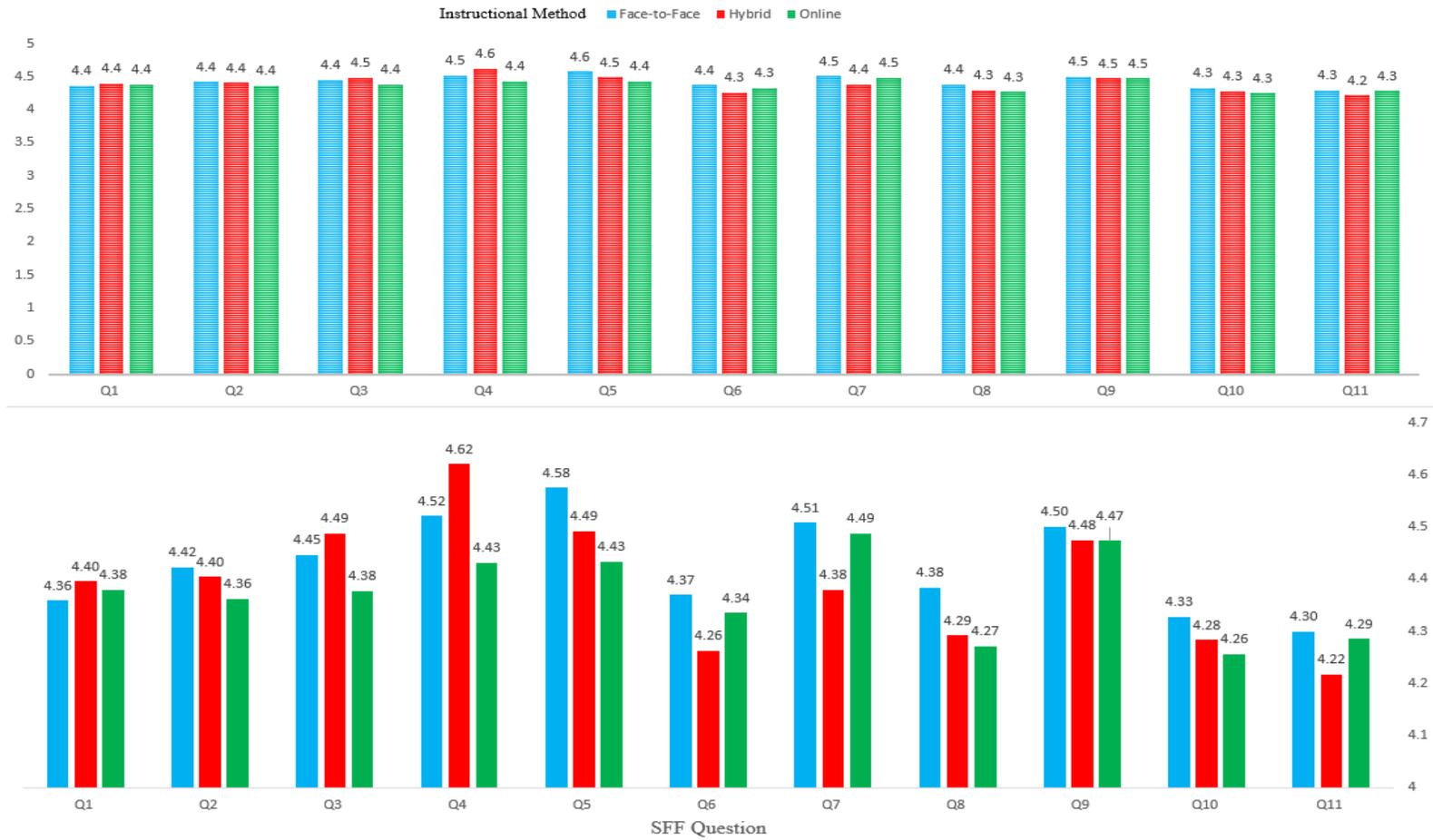Figure 4.4 Graph of SFF Means for Graduate Courses

Figure 4.5 Graph of SFF Means for Undergraduate Courses

Consideration of Student Characteristics

> *Research Question #4*: Are there significant differences in student characteristics of sex, race/ethnicity, and residency between students who enroll in face-to-face as compared to online or hybrid courses?

I derived the data to address this question from student demographic data included in each student registration record. To analyze these data, I conducted several different tests to consider a few different student characteristics. The characteristics selected included those of import to the College administration as they consider expanding their online course offerings with the goal of increasing enrollments while retaining the institutionally valued student diversity. The characteristics examined included:

- Sex
- Race / Ethnicity
- In-state students (residents) v out-of-state and international students (non-residents)

Because of the disparity – with a much higher number of face-to-face courses in the data set – to answer this question, I combined online and hybrid courses into a single group and designated that as "online." As defined by Hole (2006), Chi-square (χ^2) is a measure of the discrepancy between a set of observed and expected frequencies. Chi-square was used to determine if there was a significant relationship between the variables. A Cramer's V post-test is reported to indicate the effect size and importance.

Student Sex

To analyze sex, I constructed a cross-tabulation of variable delivery method by sex. The results, presented in Table 4.24, were significant with a very small relationship ($V = .023$) and show a slightly higher percentage of males enroll in online courses, either

Table 4.24

Sex of Students Registered in Online and Face-to-Face

Sex	Percent Face-to-Face	Percent Online
Female	87.3%	12.7%
Male	85.6%	14.4%

Chi Square = 17.26, p = .000, Cramer's V = .023

completely online or hybrid, as compared to females. This is perhaps more interesting given that females represent a super majority (69.7%) of the entire sample considered.

Student Race /Ethnicity

To analyze whether there is a difference in race, I eliminated the groups with too small sample sizes within the data set, including American Indians, Pacific Islanders, Multiracial and Unknown.

As shown in Tables 4.25 race/ethnicity has a significant effect with a small relationship. The general pattern is that Asian and International students enroll less frequently in online or hybrid courses than do White, African American or Hispanic students.

Table 4.25

Race/Ethnicity of Students Registered in Online and Face-to-Face

Race/Ethnicity	Percent Face-to-Face	Percent Online	Overall
White	86.8%	13.2%	64.96%
African American	85.8%	14.2%	16.89%
Asian	91.3%	8.7%	7.40%
Hispanic	87.0%	13.0%	7.12%
International	92.4%	7.6%	3.64%

Chi Square = 75.971, p = .000, Cramer's V = .049

Residency

Table 4.26 demonstrates that significantly more in-state residents participated in online learning than out-of-state residents. This reflects lower participation of international students, who are also coded as non-residents.

Table 4.26			
Distribution of PA and Non-PA Residents in Online and Face-to-Face			
Residency Status	Percent Face-to-Face	Percent Online	Overall
Pennsylvania Resident	85.8%	14.2%	78.37%
Non-Pennslyvania Resident	90.2%	9.8%	21.63%
Chi Square = 93.37, p = .000, Cramer's V = .053			

Summary of Analysis

As detailed above, in order compare measures of student success, persistence, and satisfaction for educational studies courses conducted with different methods of instructional delivery, I constructed a data set including courses – and student registrations within those courses – offered by a College of Education during six semesters from 2017 to 2019. I merged the course data with student responses collected via a student feedback form used across the University I eliminated from consideration course sections of dissimiliar construct to create three broad groups of comparison: face-to-face, hybrid, and online. I then conducted a series of analyses to determine the difference between these groups.

First, I examined the difference in students' academic success, as measured by course grades, and determined the interaction was statistically significant with a very

small effect. To consider the potential impact of the variation in the students' academic preparedness, I conducted a subsequent analysis of the differences in students' academic success while adjusting for cumulative GPA (as a measure of prior academic success). While this analysis showed some greater differences between instructional delivery formats for graduate students, the effect size indicated the difference was not meaningful.

Second, I examined the difference in student persisistence, as measured by course completion. As with the previous analysis, I identified a significant difference with a very small effect. However, I noted a greater variance of means between graduate and undergraduate courses and a greater difference between instructional methods for undergraduate courses.

Third, I examined the level of difference between instructional delivery methods for student satisfaction, as measured by the means reported at the course section level for student responses to the 11 questions included in the standardized University student feedback form. Although some difference was found on one question related to the quality of instructor feedback, when analyzed at the course level, this difference was not significant.

Finally, I looked at student characteristics including sex, race/ethnicity, and residency to determine if there were any notable differences in the make-up of students who registered for face-to-face versus hybrid or online courses. More differences were identified in this portion of the analyses than those conducted to address the other research questions. I determined that more males than females registered for hybrid or online courses, even though females represented a much larger percentage within the sample. I also found that Asian and International students tended to have fewer registrations in hybrid or online courses. Additionally, I reported that fewer out-of-state residents registered for hybrid or online courses.

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

As institutions increase the availability of different instructional delivery methods, including an increase in hybrid and online learning, it is important to understand the potential impact on student success, persistence, and satisfaction. Previous research has focused on aspects of instructional design (Novak, Razzouk, & Johnson, 2012) or conducted broad samples inclusive of different types of courses from a range of fields of study (Bernard et al., 2009; Gunawardena & McIsaac, 2004). This study specifically addressed the identified need for research comparing courses with similar curriculum offered (Figlio, Rush & Yin, 2010) by focusing on courses offered within the umbrella of educational studies. For courses administered by a single College of Education, this study considers how course delivery methods impacts student outcomes and student satisfaction. This chapter contains discussion and future research possibilities to help answer the research questions:

> RQ1: Is there a significant difference in student academic success, as measured by course grades, between Education courses offered through different instructional methods (online/hybrid/face-to-face)?

> RQ1A: If a difference in grades is found, does this difference remain if a student's incoming student GPA is held constant?

> RQ2: Is there a significant difference in measures of course completion between Education courses offered through different instructional methods (face-to-face/hybrid/online)?

> RQ3: Is there a significant difference in reported student satisfaction with either the course materials or instruction for Education courses offered through different instructional methods (online/face-to-face) as indicated by a Course level analysis of SFFs?

> RQ4: Are there significant differences in student characteristics between students who enroll in a face-to-face course as compared to an online or hybrid course?

In Chapter 4, the data were anlayzed around the students enrolled in courses offered through different instructional methods to consider how they performed in the classroom, if they completed the course, and how they evaluated the course Several of the findings presented are consistent with the current literature reviewed in Chapter 2. Some findings provide opportunity for future research and analysis. The discussion reviews the findings for each question and applicable literature.

Instructional Method and Student Success

RQ1: Is there a significant difference in student academic success, as measured by course grades, between Education courses offered through different instructional methods (online/hybrid/face-to-face)?

RQ1A: If a difference in grades is found, does this difference remain if a student's incoming student GPA is held constant?

The hypothesis that there will be no significant differences in academic success, as measured by final course grade, between the face-to-face, hybrid and online Education courses was rejected. However, the effect was so small that the difference cannot be considered meaningful. These findings support the findings of previous research (Means et al., 2010; Russell, 1999; Vernon-Dotson et al., 2014), which showed no meaningful difference in measures of student learning outcomes between online and face-to-face instructional modes.

To address the possible effect of differences between incoming level of preparedness of the students, and additional analysis was conducted to account for incoming GPA, as measured by the University calculated cumulative GPA for each registered student. Unlike previous studies, like Dutton and Dutton (2005), which found that online students tended to have a higher Undergraduate GPA in online versus face-to-face courses, the results of this analysis remained essentially the same. Although a degree

of additional difference was noted, the smaller effect size indicated this was not a meaningful difference.

Instructional Method and Course Completion

RQ2: Is there a significant difference in measures of course completion between Education courses offered through different instructional methods (face-to-face/hybrid/online)?

With the increase in online course enrollments (Allen & Seaman, 2013), student retention is still vitally important to the economic viability of many schools and colleges. Several studies have looked at student perisistence and course completion in online versus traditional course delivery (Johnson-Lutz et al., 2015; Lockee, Moore, & Burton, 2001; Liu, 2013) with mixed results influenced by the type of institution, characteristics of the student, and variations in financial support structures.   For this study, the hypothesis that there will be no significant differences in measures of course completion between the face-to-face, hybrid, and online Education courses was rejected, but, as before, with a very small effect size. More difference was noted between instructional delivery methods for undergraduate courses than with graduate courses. However, as with measures of student success, the effect size was so small that the difference was not meaningful.

Instructional Method and Student Satisfaction

RQ3: Is there a significant difference in reported student satisfaction with either the course materials or instruction for Education courses offered through different instructional methods (online/face-to-face) as indicated by a course level analysis of SFFs?

Chapter 2's discussion of the work of Clayson (2009), Prosser and Trigwell (1991), and Stehle, Spimath, and Kadmon (2012) summarizes the research affirming the

correlation between student evaluations of teaching (SETs) and the quality of student learning, in particular for courses with similar learning criteria and measures of student learning. In order to compare more similar groups, I included only similarly structured courses offered under the umbrella of educational studies. Additionally, in consideration of the literature, like the study by Candfield et al. (2015), that affirmed a stronger correlation between measures of learning for similar courses, I excluded from consideration courses with dissimiliar learning outcomes – including courses primarily made-up of practical application of learning in field experiences, independent research or study, and internships or other, similar, on-the-job course curriculum.

The hypothesis that there will be no significant differences in student satisfaction between the face-to-face, hybrid, and online sections of educational studies was, as for the first two research questions, rejected with a very small effect size. As shown in the results, students in these courses provided similar feedback to face-to-face, hybrid, and online instructors with only some minor differences regarding feedback received from the instructor in hybrid courses. When these differences were further examined at the graduate and undergraduate course level, no significant differences were found. Thus, to the extent that SETs can be understood as both a general measure of student satisfaction and the quality of student learning, no meaningful differences were found between education courses offered in either face-to-face, hybrid, or online formats.

Consideration of Student Characteristics

RQ4: Are there significant differences in student characteristics of sex, race/ethnicity, and residency between students who enroll in face-to-face as compared to online or hybrid courses?

When students with academic and social characteristics that align well with distance learning self-select into those instructional formats, students will succeed at least

as well as students in face-to-face classrooms. Understanding whether student characteristics are similar across delivery modes is, therefore, important when considering differences in achievement and satisfaction   As noted by Williams (2006), age, motivation, and self-discipline have previously been established as positively correlated with successful online learning.   Additionally, as discussed by Margoniner (2014), where online students tended to be older and more advanced in their degree programs, the differences between instructional delivery format did not impact the learning outcomes of students. Given the steady increase in online courses (U.S. Department of Education, 2016), Blackmon and Major (2012) and Beck and Milligan (2013) noted the general trend toward a flattening of the demographic differences between students participating in online as opposed to traditional classroom learning and suggested that, as a result, there might exist a greater effect from other academic and social characteristics when comparing student outcomes. Therefore, analysis to identify any persistent and notable demographic differences continues to provide valuable insight for administrators and instructors.

Students at the University are offered the freedom to enroll in traditional face-to-face courses, fully online courses, or blended courses that include both required synchronous or in-person meetings along with asynchronous online instruction. Students self-select into the couse section they enroll in and the population, therefore, is not randomly assigned. The reason why a student selects one delivery format over another – whether it be for preference, convenience, or scheduling reasons – is not tracked by the University. It is important, then, to consider if courses offered in a different format are more likely to exclude particular demographic groups of students from participation. By addressing barriers to access based on demographic student characteristics, policy-makers and instructors may address potential differences in academic success and persistence otherwise attributed to instructional delivery method.

Examining the demographics of students in this sample, my study considered the relationship between the selected instructional delivery method and key characteristics of the students. To account for the disparity in group size, for this analysis, I combined student registrations for hybrid and online learning into a single group and compared it to the whole of the face-to-face group. The two-group sample was of sufficient size to detect at least a small effect for: race/ethnicity, and student residency status. It was found that more males than females enrolled in online or hybrid courses, but the relationship between instructional method and student sex was very small (V=.023). Asian students were somewhat less likely to enroll in online or hybrid courses. While, overall, the majority of students participate in face-to-face courses at this institution, in-state residents, as compared to out-of-state residents (inclusive of international students), are more likely to register for hybrid and online courses.

As noted previously, the University values student diversity and references its "diverse community of learners" in its mission statement (2020). Therefore, it is perhaps important for the College administration and its instructors to observe that this study demonstrated notable differences in the characteristics of the students who enrolled in hybrid or online learning. Given that Asian students represent the third-highest category of race/ethnicity for the University (after Whites and African Americans), and a majority of female students make up the current student enrollment in education studies, it may be useful to further investigate the intersection of race, ethnicity, and sex in student selection of online or hybrid learning opportunities.

Summary of Findings

Taken together, the findings for the first three research questions are in line with the "no meaningful significant difference" findings of the existing literature on

comparisions of differentiated instructional method for courses across disciplines. Substantial differences in student success, course completion, and student feedback were not observed, indicating similar learning outcomes for students regardless of delivery format. For the College of Education, it is important for accreditation, and also for its own goals, to ensure that all students are provided with high quality instruction across all delivery modes. The results from this study confirm that student achievement is equitable in the online, hybrid, and face-to-face delivery modes. These findings should reassure administrators and instructors seeking to increase the availability of distance learning opportunities for students in educational studies that do not include practical or experiential learning.

At the same time, differences in the characteristics of students who enroll in hybrid and online courses, as opposed to courses offered through traditional classroom instruction, indicate the presence of some potential barriers to these instructional methods for educational studies students. Men more frequently enrolled in online or hybrid courses than women, even though women make up a larger majority of education studies students. Additionally, Asian students, which represent the second-largest minority group in the College, chose hybrid or online learning less frequently than other students. The College will need to consider how any planned changes in course offerings may impact these student demographics. And, if the College, as part of a state public institution, sees distance learning as an opportunity to expand its student population beyond the boarders of the state, the perceived lower enrollments of out-of-state residents into hybrid and online courses may negatively impact these goals. Before dramatically increasing its course offerings, I recommend the College conduct additional research into the motivations and barriers for these groups of students that might discourage enrollment into hybrid and online learning courses.

Limitations of the Study and Recommendations for Future Research

All data were collected prior to the beginning of this study, using existing databases and instruments. I acknowledge that other more extensive methods could have been used. The data gathered and analyzed here were prioritized at the request of the administration in the College to advance their understanding of the potential impact of increasing online and hybrid instruction across programs and in the hope that the study results could positively impact future course offerings. As a function of the historical preference for the College to offer courses primarily in traditional face-to-face classrooms, the sample included much smaller groups of course sections offered in either hybrid or online formats. This pre-existing difference between the groups impacts the noted effects and generalizability of the study results.

I should also note that, in my role in the University, I support the technology acquisition and implementation for the College of Education, including providing direct support to individual instructors who elect to offer a course in hybrid or online format. However, this support represents only a portion of my responsibilities and I have no direct influence on selecting or appointing instructors or developing course content. Other staff members, such as department chairs and the director of academic operations, determine which courses or sections are offered using which instructional method.

Of additional note, this study did not include analysis that considered the potential influence of the variation in the type of instructor assigned to courses offered in different instructional methods. Potentially, some instructors taught using only one type of instructional method and some instructors taught the same course using different types of instructional method. Further exploration of the variation in instructor seniority and rank – as well as other characteristics of the instructor such as sex and race/ethnicity – within

and between the groups may provide a more comprehensive understanding of the inter-relationship between instructor, student, and instructional method.

While differences in student feedback were not identified in this study, it is worth noting that the impact of characteristics of the instructor (e.g., age, sex, seniority) were also not considered. The potential for bias in factors of instructor demographics remains unclear from the current literature. Additionally, my analysis of student characteristics included only a limited set of possible variables (sex, race/ethnicity, and residency status). Expanded data would provide opportunity to conduct further analysis that includes additional descriptive categories for students, such as age, status of degree completion, and specific area of study within educational studies. Further research addressing these factors for both instructors and students may provide a more nuanced understanding of these findings.

The structure of the study, which used only existing data, did not allow me to conduct a fully randomized experiment that controlled for specific course content or instructor characteristics: both of which may impact the validity, reliability, and generalizability of the results. Although the response rate was generally higher for educational studies courses, the course evaluation satisfaction data were only available at the course section level, rather than the student level. Differences in the sample size among the groups, a limitation of the available data, also present a threat to validity. Further research comparing more similar groups may provide additional insight, as would analysis of the open-ended response questions included in the student feedback form. These qualitative data were not available for my study, but may provide greater insight into the satisfaction of students across different modalities of teaching. Additionally, as a fundamental aspect of educational studies in HE includes the preparation of teachers and administrators for professional careers in teaching and learning, future research should consider the more complex learning environments represented by field experiences,

practicums, and internships that are a common element of most education degree programs.

It was also noted that the completion rate of course evaluations in the sample was generally higher than the completion rate for the University. A higher response rate for student feedback forms in the College of Education may be reflective of the college's efforts to invite and solicit student feedback. It may also be reflective of the College's focus on issues of assessment in education, which may make education students more likely to engage in evaluations of courses. Additional research into the motivations and participation of students in course evaluations may be beneficial to understanding these differences. However, it was also noted that the completed feedback forms for online courses were notably lower than the other two comparison groups. In particular, the response rate for undergraduate online courses was the lowest at 44% -- a response rate more in-line with the overall University. Given that the University uses the same method to collect student feedback forms, regardless of the course delivery format, it seems worthy of further study to explore why students who would be assumed to be more frequently engaged in other online activities for the support of their online learning were also less likely to complete online evaluations of their courses.

As I also reported, more in-state residents versus out-of-state residents and international students registered for hybrid and online courses. While this study does not seek to address causation, it is worth considering that reductions in enrollments for out-of-state and international students may be a function of the regulatory limitations on public institutions of higher education, which must seek additional permissions from relevant accrediting bodies before offering courses to students who reside outside the regional area of the (U.S. Department of Education, 2016). Further investigation into the barriers to enrollment for students and, more broadly, the role of policy on student participation in online learning is warranted.

Conclusion

The results of this study are important because they add to the extensive literature supporting "no meaningful significant difference" on student achievement, persistence, and satisfaction between face-to-face, hybrid, and online instructional methods. In particular, this study extends this research by specifically focusing on courses of similar curriculum as offered within educational studies. The results reinforce that administrators and instructors should continue to expand access to courses through the flexibility of online and hybrid learning. However, as programs expand their course offerings, further investigation is warranted into the barriers to hybrid and online learning for certain groups of educational studies students within this institution, including women, Asians, and out-of-state residents.

In consideration of the differences identified here, this study revealed several additional opportunities for further research. Given the large percentage of courses offered by the College that included some aspect of experiential learning (>11%), analysis focused on the relationship between instructional method and measures of student success that included this additional facet of learning in educational studies would provide valuable data for administrators and policy-makers. Qualitative analysis of the open-ended responses to student evaluations of teaching may afford additional insight into the differences in student satisfaction with hybrid and online courses. Data inclusive of additional characteristics of both the assigned instructor (e.g., sex, race/ethnicity, seniority/tenure status) and the students (e.g., age, degree program, time to completion) would support a multivariate regression method of analysis to further identify and understand factors in the relationships between and among student participation in different instructional modalities. Leveraging these recommendations for further research while managing the limitations noted could enhance the understanding of the potential

implications for student participation and outcomes, as higher education continues the long-term trend toward increased offerings of hybrid and online courses.

REFERENCES CITED

Allen, E. & Seaman, J. (2010). *Class differences: Online education in the United States, 2010.* Babson Survey Research Group. Sloan Consortium. http://files.eric.ed.gov/fulltext/ED529952.pdf

Allen, E. & Seaman, J. (2013). *Changing course: Ten years of tracking online education in the United States*. Babson Survey Research Group. Sloan Consortium. https://files.eric.ed.gov/fulltext/ED541571.pdf

Allen, I. E., & Seaman, J. (2015). *Grade level: Tracking online education in the United States*. The Babson Survey Research Group, OLC, Pearson and Tyton Partners. http://www.onlinelearningsurvey.com/reports/gradelevel.pdf

Alexander, Susan M., and Sonalini Sapra. (2013) "'Post it on the wall!': Using Facebook to complement student learning in gender and women's studies courses." *Feminist Teacher*, *23*(2), p. 142+. https://link-gale-com.libproxy.temple.edu/apps/doc/A393518490/AONE?u=temple_main&sid=AONE&xid=f35b0341.

Alter, M., & Reback, R. (2014). True for your school? How changing reputations alter demand for selective U.S. colleges. *Educational Evaluation and Policy Analysis*, *36*(3), 346–370. doi: 10.3102/0162373713517934

Anderson, L. & Krichbaum, K.E. (2017). Best practices for learning physiology: Combining classroom and online methods. *Advances in Physiology Education*, 05 July 2017. doi: 10.1152/advan.00099.2016

Baldwin, T., & Blattner, N. (2003). Guarding against potential bias in student evaluations. *College Teaching, 51*(1), 27. http://libproxy.temple.edu/login?url=https://search-proquest-com.libproxy.temple.edu/docview/274595320?accountid=14270

Bartlett, S. & Burton, D. (2007) *Introduction to education studies* (2nd ed). Sage Publications.

Bathke, A.C., Schabenberger, O., Randall, D.T., & Madden, L. (2009). Greenhouse-Geisser adjustment and the ANOVA-type statistic: cousins or twins? *The American Statistician*, *63*(3), 239-246. doi: 10.1198/tast.2009.08187

Beck, H. & Milligan, M. (2013). Factors influencing the institutional commitment of online students. *Internet and Higher Education. 20*, pp. 51–56.

Becker, W. H., Bosshardt, W., & Wattas, M. (2012). How departments of economics evaluate teaching. *The Journal of Economic Education, 43*(3), 325-333. doi: 10.1080/00220485.2012.686826

Bernard, R., Abrami, P., Borokhovski, E., Wade, C. A., Tamim, R., Surkes, M., & Bethel, E. (2009). A meta-analysis of three types of interaction treatments in distance education. *Review of Educational Research*, *79*(3), 1243-1289.

Blackmon, S. & Major, C. (2012). Student experiences in online courses: A qualitative research synthesis. *The Quarterly Review of Distance Education. 13*(2), pp. 77-85.

Boehmer, D.M. & Wood, W.C., (2017) Student vs. faculty perspectives on quality instruction: Gender bias, "hotness," and "easiness" in evaluating teaching, *Journal of Education for Business, 92*(4), 173-178. doi: 10.1080/08832323.2017.1313189

Bond, L. (2007). The case for common examinations. *The Hispanic Outlook in Higher Education, 17*(14), 31.

Bonvllian, W. B., & Singer, S. R. (2013). The online challenge to higher education. *Issues in Science and Technology, 29*(4), 23-30. http://search.proquest.com/docview/1506956003?accountid=14270

Bowen, W. (2013). *Higher education in the digital age*. Princeton University Press: Princeton, NJ.

Bowen, W., Chingos, M., Lack, K. & Nygrn, T. (2013). Online learning in higher education. *Education Next*, *13*(2), 59-64. http://educationnext.org/files/ednext_XIII_2_bowen_chingos.pdf

Brinthaupt, T.M., Fisher, L.S., Gardner, J.G., Raffo, D.M., & Woodard, J.B. (2011). What the best online teachers should do. *Journal of Online Learning and Teaching*, *7*(4), 515. http://search.proquest.com/docview/1497202433?accountid=14270

Burns, D. J. (2011). How can students improve their performance? An examination of possible correlates. *Business Education Digest*, *18*, 1-10.

Canfield, M. L., Kivisalu, T. M., Van Der Karr, C., King, C., & Phillips, C. E. (2015). The use of course grades in the assessment of student learning outcomes for general education. *SAGE Open*. doi: 10.1177/2158244015615921

Centra, J.A. & Gaubatz, N.B. (2000) Is there gender bias in student evaluatons of teaching? *The Journal of Higher Education*, *71*(1), 17-33. doi: 10.1080/00221546.2000.11780814

Christensen, C. (2011). *Disrupting class: How disruptive innovation will change the way the world learns*. New York: McGraw Hill.

Clayson, D. E. (2009). Student evaluations of teaching: Are they related to what students learn? A metaanalysis and review of the literature. *Journal of Marketing Education*, *31*, 16–30.

Cochran, J. D., Campbell, S. M., Baker, H. M., & Leeds, E. M. (2014). The role of student characteristics in predicting retention in online courses. *Research in Higher Education, 55*(1), 27-48. doi: 10.1007/s11162-013-9305-8

Cohen, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. *Review of Educational Research*, *51*, 281–309.

Crawford-Ferre, H. G., & Wiest, L. R. (2012). Effective online instruction in higher education. *Quarterly Review of Distance Education, 13*(1), 11.

Dalelio, C. (2013). Student participation in online discussion boards in a higher education setting. *International Journal on E-Learning, 12*(3), 249.

Demirci, N. (2014). What is massive open online courses (MOOCs) and what is promising us for learning?: A review-evaluative article about MOOCs. *Necatibey Faculty of Education, 8*, 231-256.

Dervarics, C. (2013). Fiscal issues provide roadblock for higher ed in 2013. *Diverse Issues in Higher Education*, 5, np.

Dhilla, S. J., (2017): The role of online faculty in supporting successful online learning enterprises: A literature review. *Higher Education Politics & Economics*, *3*(1). https://temple.userservices.exlibrisgroup.com/view/action/uresolver.do?operation=resolveService&package_service_id=18466243930003811&institutionId=3811&customerId=3810

Downing, J. J., & Dyment, J. E. (2013). Teacher educators' readiness, preparation, and perceptions of preparing preservice teachers in a fully online environment: An exploratory study. *Teacher Educator, 48*(2), 96.

Dutton, J. & Dutton, M. (2005) Characteristics and performance of students in an online section of business statistics. *Journal of Statistics Education*, *13*(3), n.p. doi: 10.1080/10691898.2005.11910564

Dutton, J., Dutton, M. & Perry, J. (2001), Do online students perform as well as lecture students? *Journal of Engineering Education*, *90*, 131–136.

Elliott, C., & Soo, K. T. (2013). The international market for MBA qualifications: The relationship between tuition fees and applications. *Economics of Education Review, 34*, 162-174. doi:10.1016/j.econedurev.2013.02.006

Evans, J. (2013). Exorbitant education: The rising cost of tuition surpasses that of inflation. *Diverse Issues in Higher Education, 29*(26), 6.

Felton, D. A. (2013). Will student debt impact our specialty? *Journal of Prosthodontics : Official Journal of the American College of Prosthodontists, 22*(6), 431-431. doi:10.1111/jopr.12087

Felton, J., J. Mitchell, and M. Stinson (2004). Web-based student evaluation of professors: The relations between perceived quality, easiness and sexiness. *Assessment and Evaluation in Higher Education*, *29*(1), 91–108.

Field, A. (2006). Cramer's V. In G. Davey, *Encyclopaedic dictionary of psychology*. Routledge. Credo Reference: http://libproxy.temple.edu/login?url=https://search.credoreference.com/content/entry/hodderdpsyc/cramer_s_v/0?institutionId=1644

Figlio, D., Rush, M., & Yin, L. (2010). *Is it live or is it internet? Experimental estimates of the effects of online instruction on student learning.* Working Paper 16089. National Bureau of Economic Research: Cambridge, MA. http://www.nber.org/papers/w16089.pdf

Forsey, M., Low, M., & Glance, D. (2013). Flipping the sociology classroom: Towards a practice of online pedagogy. *Journal of Sociology, 49*(4), 471-485. doi:10.1177/1440783313504059

Ginder, S.A., Kelly-Reid, J.E., and Mann, F.B. (2018). *Enrollment and employees in postsecondary institutions, fall 2017; and financial statistics and academic libraries, fiscal year 2017: First look (provisional data).* (NCES 2019- 021rev). U.S. Department of Education. National Center for Education Statistics: Washington, D.C. http://nces.ed.gov/pubsearch.

Grummon, P. T. H. (2010). Trends in higher education. *Planning for Higher Education, 39*(1), 51.

Gunawardena, C., & McIssac, M. (2004). Distance education. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (2nd ed., pp. 355-396). Erlbaum. https://psycnet.apa.org/record/2004-00176-000

Guthrie, K. M. (2012). Barriers to the adoption of online learning systems. *EDUCAUSE Review, 47*(4), 50.

Haywood, J.L., & Murty, K. S. (2018). Undergraduate student perceptions and satisfaction toward online psychology courses at an HBCU. *Race, Gender & Class, 25*(1), 131-146. http://libproxy.temple.edu/login?url=https://search-proquest-com.libproxy.temple.edu/docview/2233853313?accountid=14270

Hole, G. (2006). Chi-square test. In G. Davey, *Encyclopaedic dictionary of psychology*. Routledge. http://libproxy.temple.edu/login?url=https://search.credoreference.com/content/entry/hodderdpsyc/chi_square_test/0?institutionId=1644

Hu, S. (2005). Beyond grade inflation: Grading problems in higher education. *ASHE Higher Education Report, 30*(6), n.p.

Ithindi, E. T. (2013). Using discussion forums to mediate learning in higher education: A literature review through an activity theory lens. *Nawa: Journal of Language and Communication, 7*(2), 1.

Johnson-Lutz, H., Sessoms-Penny, S., Schneider, J., & Underdahl, L. (2015). Best practices relating to student persistence, motivation, and retention in online MBA programs. *The Exchange, 4*(2), 25-32. http://libproxy.temple.edu/login?url=https://search-proquest-com.libproxy.temple.edu/docview/1818006681?accountid=14270

Jones, J., Gaffney-Rhys, R & Jones, E. (2014) Handle with care! An exploration of the potential risks associated with the publication and summative usage of student evaluation of teaching (SET) results, *Journal of Further and Higher Education*, *38*(1), 37-56. doi: 10.1080/0309877X.2012.699514

Kahlenberg, R. D. (2013, May). Profs in the cloud. *The Washington Monthly, 45*, 60-61. http://search.proquest.com/docview/1353347605?accountid=14270

Kamenetz, A. (2013). From baumol's cost disease to moore's law: Bending the cost curve in higher education. In A. Kelly & K. Carey (Eds.), *Stretching the higher education dollar: How innovation can improve access, equity, and affordability* (pp. 9-26). Harvard Education Press. https://www.gse.harvard.edu/news/13/10/harvard-edcast-stretching-higher-education-dollar

Keengwe, J., & Kang, J. (2012). Blended learning in teacher preparation programs: A literature review. *International Journal of Information and Communication Technology Education*, 8(2), 81-93. doi:10.4018/jicte.2012040107

Keengwe, J., & Kidd, T. T. (2010). Towards best practices in online learning and teaching in higher education. *Journal of Online Learning and Teaching, 6*(2), 533. http://search.proquest.com/docview/1497197857?accountid=14270

Keramidas, C. G. (2012). Are undergraduate students ready for online learning? A comparison of online and face-to-face sections of a course. *Rural Special Education Quarterly*, *31*(4), 25.

Kogan, J. 2014. Student course evaluation: Class size, class level, discipline, and gender bias. Paper presented at the sixth International Conference on Computer

Supported Education, April, Barcelona, Spain.
https://files.ifi.uzh.ch/stiller/CLOSER%202014/CSEDU/CSEDU/Learning/Teaching%20Methodologies%20and%20Assessment/Short%20Papers/CSEDU_2014_140_CR.pdf

Klaus, T., & Changchit, C. (2014). Environmental factors of distance learning: An exploratory study. *International Journal of Information and Communication Technology Education, 10*(1), 14.

Kirtman, L. (2009). Online versus in-class courses: An examination of differences in learning outcomes. *Issues in Teacher Education, 18*(2), 103-116. http://libproxy.temple.edu/login?url=https://search-proquest-com.libproxy.temple.edu/docview/233320851?accountid=14270

Kogan, L.R., Schoenfeld-Tacher, R.& Hellyer, P.W. (2010). Student evaluations of teaching: Perceptions of faculty based on gender, position, and rank. *Teaching in Higher Education, 15*(6), 623-636.

Legutko, R. S. (2007). Face-to-face or cyberspace: Analysis of course delivery in a graduate educational research course. *Journal of Online Learning and Teaching, 3*(3), n.p.. http://jolt.merlot.org/vol3no3/legutko.htm

Liebe, K. (2013). Audiology students give a view from the ground on rising tuitions. *The Hearing Journal, 66*(2), 28.

Liu, M. (2013). Disrupting teacher education: High costs for brick-and-mortar degrees create opportunities for online programs. *Education Next, 13*(3), 26.

Lockee, B., Moore, M., & Burton, J. (2001). Old concerns with new distance education research. *Educause Quarterly*, 60-63.

Lowenthal, P., Bauer, C., & Chen, K. (2015) Student perceptions of online learning: An analysis of online course evaluations. *American Journal of Distance Education, 29*(2), 85-97. doi: 10.1080/08923647.2015.1023621

Luzer, D. (2013, May). Revolution for thee, not me. *The Washington Monthly, 45*, 58-59. http://search.proquest.com/docview/1353347600?accountid=14270

Manning, K. (2013). New science. In *Organizational theory in higher education* (pp. 135-149). Routledge. https://books.google.com/books/about/Organizational_Theory_in_Higher_Educatio.html?id=EOqoLw06OWMC

Margoniner, V. (2014). Learning gains in introductory astronomy: Online can be as good as face-to-face. *The Physics Teacher*, *52*(5), 298.

McCrory, R., Putnam, R. & Jansen, A. (2008). Interaction in online courses for teacher education: Subject matter and pedagogy. *Journal of Technology and Teacher Education*, *16*(2), 155-180. https://search-proquest-com.libproxy.temple.edu/docview/200001834/fulltextPDF/9DC0CFCD48194500PQ/1?accountid=14270

McGorry, S. Y. (2012). No significant difference in service learning online. *Journal of Asynchronous Learning Networks, 16*(4), 45.

McNatt, D. (2019). Service learning. *Education & Training*, *61*(2), 129-144.

McNutt, M. (2013). Bricks and MOOCs. *Science (New York, N.Y.), 342*(6157), 402-403. doi:10.1126/science.1246943

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, J. (2009). *Evidence of evaluation 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. https://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf

Moody's Investors Service (2010). *Special comment: Annual sector outlook for U.S. higher education for 2010* (Report Number: 122463). Mount Holyoke. https://www.mtholyoke.edu/sites/default/files/iplanning/docs/annualsectoroutlookforushighereducation2010.pdf

Muljana, P. S., & Tian Luo. (2019). Factors contributing to student retention in online learning and recommended strategies for improvement: A systematic literature review. *Journal of Information Technology Education*. *18*, 19–43. doi: 10.28945/4182

Murdock, J., Williams, A., Becker, K., Bruce, M. A., & Young, S. (2012). Online versus on-campus: A comparison study of counseling skills courses. *The Journal of Human Resource and Adult Learning, 8*(1), 105.

Novak, E., Razzouk, R., & Johnson, T. (2012). The educational use of social annotation tools in higher education: A literature review. *Internet and Higher Education, 15*(1), 39-49. doi:10.1016/j.iheduc.2011.09.002

Otter,R., Seipel, S., Graeff, T., Alexander, B., Boaraiko, C., Gray, J., Petersen,K., & Sadler, K. (2013). Comparing student and faculty perceptions of online and traditional courses. *Internet and Higher Education. 19*(2013), 27-35.

Palmer, M. M., Shaker, G., & Hoffmann-Longtin, K. (2014). Despite faculty skepticism: Lessons from a graduate-level seminar in a hybrid course environment. *College Teaching, 62*(3), 100.

Patterson, B. & McFadden, C. (2009). Attrition in online and campus degree programs. *Online Journal of Distance Learning Administration*, *2*(2), n.p.

Peterson, D., Biederman, L. A., Andersen, D., Ditonto, T. M., & Roe, K. (2019). Mitigating gender bias in student evaluations of teaching. *PloS one*, *14*(5), e0216241. doi: 10.1371/journal.pone.0216241

Pløhn, T. (2014). Pervasive learning--using games to tear down the classroom walls. *Electronic Journal of e-Learning, 12*(3), 299-311. http://libproxy.temple.edu/login?url=https://search-proquest-com.libproxy.temple.edu/docview/1651861215?accountid=14270

Rafferty, P. D., & Anderson, C. M. (2013). Are executive students satisfied with online-learning experiences? implications for online training and executive education. *Training & Management Development Methods, 27*(4), 673.

Rakap, S., Jones, H., & Emery, A.K. (2014). Evaluation of a web-based professional development program (project ace) for teachers of children with autism spectrum disorders. *Teacher Education and Special Education,* pp. 1-19. doi: 10.1177/0888406414535821

Roberts, J., Crittenden, L., & Crittenden, J. (2011). Students with disabilities and online learning: A cross-institutional study of perceived satisfaction with accessibility compliance and services. *Internet and Higher Education, 14*(4), 242-250. doi:10.1016/j.iheduc.2011.05.004

Sandeen, C. (2013). Integrating MOOCS into traditional higher education: The emerging "MOOC 3.0" era. *Change: The Magazine of Higher Learning, 45*(6), 34-39. doi:10.1080/00091383.2013.842103

Schrum, K., & Sleeter, N. (2013). Teaching history online: Challenges and opportunities. *Magazine of History*, *27*(3), 35.

Schneider, J. & Hutt, E. (2014) Making the grade: a history of the A–F marking scheme, *Journal of Curriculum Studies*, 46(2), 201-224. doi: 10.1080/00220272.2013.790480

Secret, M., Bentley, K.J., & Kadolph, J.C., (2016) Student voices speak quality assurance: Continual improvement in online social work education. *Journal of Social Work Education,* 52:1, 30-42. doi: 10.1080/10437797.2016.1112630

Seaman, J.E., Allen, I.E., & Seaman, J. (2018). *Grade increase: Tracking distance education in the United States*. Babson Survey Research Group. https://onlinelearningsurvey.com/reports/gradeincrease.pdf

Selingo, J. (2013). *College unbound: The future of higher education and what it means for students*. Houghton Mifflin Harcourt Publishing. https://jeffselingo.com/book/college-un-bound/

Selingo, J. (2013, 6 December). Colleges Can Still Save Themselves. Here's How. *The Chronicle of Higher Education*, *60*(14). http://go.galegroup.com/ps/i.do?id=GALE%7CA364442265&v=2.1&u=temple_main&it=r&p=AONE&sw=w&asid=ad4e320c572dc3c7828c3ab6f0d58a45

Singh, V. (2013). Role of ict in higher education: An overview. *Indian Streams Research Journal,* 2277.

Smith, G. G., Sorensen, C., Gump, A., Heindel, A. J., Caris, M., & Martinez, C. D. (2011). Overcoming student resistance to group work: Online versus face-to-face.

The Internet and Higher Education, 14(2), 121-128. doi:10.1016/j.iheduc.2010.09.005

Stehle, S., Spinath, B., & Kadmon, M. (2012). Measuring teaching effectiveness: Correspondence between students' evaluations of teaching and different measures of student learning. *Research in Higher Education*, *53*(8), 888-904. doi:10.1007/s11162-012-9260-9

Stone-Brown, K. (2013). Bleak-financial outlook faced by medics. *Student BMJ*, *21*, 10.

Suskie, L. A. (2009*). Assessing student learning: A common sense guide*. Jossey-Bass. https://www.wiley.com/en-us/Assessing+Student+Learning%3A+A+Common+Sense+Guide%2C+3rd+Edition-p-9781119426936

Tabachnick, B. G., & Fidell, L. S. (2000). *Using multivariate statistics*. Allyn and Bacon/Pearon Educaton. https://psycnet.apa.org/record/2006-03883-000

Temple University (2020). e-SFF. https://esff.temple.edu

Temple University (2020). *Temple University Fact Book: 2019-2020*. https://www.temple.edu/ira/documents/data-analysis/Fact-Book/TU_Fact_Book_2019-2020.pdf

Thompson, J. R., Klass, P. H., & Fulk, B. M. (2012). Comparing online and face-to-face presentation of course content in an introductory special education course. *Teacher Education and Special Education*, *35*(3), 228-242. doi:10.1177/0888406412441627

Trigwell, K., & Prosser, M. (1991). Improving the quality of student learning: the influence of learning context and student approaches to learning on learning outcomes. *High Educ* 22, 251–266. doi: 10.1007/BF00132290

U.S. Department of Education/National Center for Education Statistics. (2016*).* Table 303.10. Total fall enrollment in degree-granting postsecondary institutions, by attendance status, sex of student, and control of institution: Selected years, 1947 through 20206. *Digest of education statistics, 2016.* (NCES 2015-16). https://nces.ed.gov/programs/digest/d16/tables/dt16_303.10.asp

U.S. Department of Education Office of Planning, Evaluation, and Policy Development
Policy and Program Studies Service (2010). *Evaluation of evidence-based
practices in online learning: A meta-analysis and review of online learning
studies*. http://www2.ed.gov/rschstat/eval/tech/evidence-based-
practices/finalreport.pdf

U.S. Government Accountability Office. (26 Sept 2002). *Growth in distance education
programs and implications for federal education policy*. (GAO-02-1125T).
http://www.gao.gov/products/GAO-02-1125T

Uttl, B., & Smibert, D. (2017). Student evaluations of teaching: Teaching quantitative
courses can be hazardous to one's career. *PeerJ.* doi:10.7717/peerj.3299

Vernon-Dotson, L. J., Floyd, L. O., Dukes, C., & Darling, S. M. (2014). Course delivery:
Keystones of effective special education teacher preparation. *Teacher Education
and Special Education*, *37*(1), 34-50.

Voss, B. D. (2013). MOOCs: Get in the game. *EDUCAUSE Review, 48*(1), 58.

Warner, A. G. (2016). Developing a Community of Inquiry in a Face-to-Face Class: How
an Online Learning Framework Can Enrich Traditional Classroom
Practice. *Journal of Management Education*, *40*(4), 432–452. doi:
10.1177/1052562916629515

Warren, L. L., & Holloman, H. L., Jr. (2005). On-line instruction: Are the outcomes the
same? *Journal of Instructional Psychology, 32*(2), 148-151.
http://libproxy.temple.edu/login?url=https://search-proquest-
com.libproxy.temple.edu/docview/62005604?accountid=14270

Weldy, T. G. (2018). Traditional, blended, or online: Business student preferences and
experience with different course formats. *E-Journal of Business Education and
Scholarship of Teaching, 12*(2), 55-62.
http://libproxy.temple.edu/login?url=https://search-proquest-
com.libproxy.temple.edu/docview/2155997410?accountid=14270

Wildavsky, B. (2013). Classes for the masses. In A. Kelly & K. Carey (Eds.), *Stretching
the higher education dollar: How innovation can improve access, equity, and
affordability* (pp. 125-144). Harvard Education Press.

Wilhelm, W. B. (2004). The relative influence of published teaching evaluations and other instructor attributes on course choice. *Journal of Marketing Education*, *26*(1), 17–30. doi: 10.1177/0273475303258276

Williams, S. L. (2006). The effectiveness of distance education in allied health science programs: A meta-analysis of outcomes. *American Journal of Distance Education*, *20*(3), 127-141. doi:10.1207/s15389286ajde2003_2

Wilson, J., Beyer, D., & Monteiro, H. (2014). Professor age affects student ratings: Halo effect for younger teachers. *College Teaching*, 62, 20-24. doi:10.1080/87567555.2013.825574

Wood, J., Kiggins, R., & Kickham, K. (2017). Food or thought? Assessing internal and external factors affecting evaluations of instructor effectiveness. *Journal of Political Science Education, 13*(4), 373-388. doi:10.1080/15512169.2017.1329089

Wright, S.L. & Jenkins-Guarnieri, M.A., (2012). Student evaluations of teaching: combining the meta-analyses and demonstrating further evidence for effective use. *Assessment & Evaluation in Higher Education, 37*(6), 683-699. doi: 10.1080/02602938.2011.563279

Zaiontz, C. (2014). Effect size for Chi-square test. *Real statistics using excel*. http://www.real-statistics.com/chi-square-and-f-distributions/effect-size-chi-square/

APPENDIX A: ANALYSIS OF SFF QUESTIONS RESPONSE RATES

An analysis of the response rates was completed with significant difference as a function of instructional method and a significant interaction between instructional method and the questions. However, the interaction between instructional method, course level, and the questions was not significant. These results are presented in here but were not considered relevant for general discussion.

Table A.1 presents the means and standard deviations for each of the 11 questions on the student feedback form. Table A.2 presents the summary of the ANOVA. There is a significant difference as a function of instructional method with a medium effect size and a significant interaction between instructional method and question with a medium effect size. No significant difference was found in the interaction between question response rate, course level, and instructional method.

As shown by Figures A.1 and A.2, there was a notable diffence by instructional method for several of the question response rates. Additional analysis was done considering only differences by only Instructional Method and not couse level, since the triple interaction was not significant. Table A.3 presents the means and standard deviations of the response rate for the 11 SFF questions when graduate and undergraduate response rates are combined into a single group. There were significant differences on six of the questions. Table A.4 presents the means and standard deviations for SFF questions one, three, four, five, seven, and eight, by instructional method with course level combined into a single group.

A review of the data showed that for question eight and question nine, the differences were primarily found in undergraduate courses and the effect sizes were very small. Additional post-hoc analysis was done for questions one, three, four, and five. Table A.5 presents the results of the Tukey analysis for each question. Except for question five, in which there was a significant difference found between the response

Table A.1

SFF Questions Response Rate by Course Level Means and Standard Deviations

SFF Question	Course Level	Instructional Method	Mean	Std. Deviation	N
Q1 Response Rate	Graduate	Face-to-Face	99.40%	2.64%	403
		Hybrid	99.02%	3.20%	105
		Online	94.77%	7.96%	61
	Undergraduate	Face-to-Face	99.41%	2.77%	952
		Hybrid	99.44%	1.88%	24
		Online	95.46%	7.18%	66
Q2 Response Rate	Graduate	Face-to-Face	99.47%	2.27%	403
		Hybrid	99.42%	2.37%	105
		Online	98.50%	5.31%	61
	Undergraduate	Face-to-Face	99.57%	2.77%	952
		Hybrid	100.00%	0.00%	24
		Online	99.86%	1.12%	66
Q3 Response Rate	Graduate	Face-to-Face	98.36%	4.22%	403
		Hybrid	97.48%	6.73%	105
		Online	96.70%	7.08%	61
	Undergraduate	Face-to-Face	98.68%	3.64%	952
		Hybrid	98.55%	3.56%	24
		Online	96.75%	5.98%	66
Q4 Response Rate	Graduate	Face-to-Face	98.70%	3.49%	403
		Hybrid	98.96%	3.42%	105
		Online	88.42%	13.70%	61
	Undergraduate	Face-to-Face	98.32%	4.34%	952
		Hybrid	99.07%	3.14%	24
		Online	87.96%	11.07%	66
Q5 Response Rate	Graduate	Face-to-Face	98.76%	3.72%	403
		Hybrid	97.35%	5.84%	105
		Online	90.36%	12.99%	61
	Undergraduate	Face-to-Face	98.48%	4.07%	952
		Hybrid	97.39%	4.96%	24
		Online	89.80%	12.36%	66

Table A.1

(continued)

SFF Question	Course Level	Instructional Method	Mean	Std. Deviation	*N*
Q6 Response Rate	Graduate	Face-to-Face	98.98%	3.33%	403
		Hybrid	99.22%	2.80%	105
		Online	99.17%	3.50%	61
	Undergraduate	Face-to-Face	99.49%	2.53%	952
		Hybrid	99.44%	1.88%	24
		Online	99.03%	3.38%	66
Q7 Response Rate	Graduate	Face-to-Face	96.34%	8.93%	403
		Hybrid	94.87%	9.27%	105
		Online	98.43%	3.72%	61
	Undergraduate	Face-to-Face	98.12%	4.62%	952
		Hybrid	98.89%	3.13%	24
		Online	98.21%	4.67%	66
Q8 Response Rate	Graduate	Face-to-Face	99.53%	2.02%	403
		Hybrid	99.46%	2.29%	105
		Online	98.64%	4.22%	61
	Undergraduate	Face-to-Face	99.59%	2.29%	952
		Hybrid	99.12%	2.37%	24
		Online	98.95%	3.98%	66
Q9 Response Rate	Graduate	Face-to-Face	98.64%	3.67%	403
		Hybrid	97.98%	6.28%	105
		Online	98.52%	6.73%	61
	Undergraduate	Face-to-Face	98.41%	5.06%	952
		Hybrid	97.81%	3.64%	24
		Online	98.29%	4.24%	66
Q10 Response Rate	Graduate	Face-to-Face	99.21%	2.58%	403
		Hybrid	99.31%	2.61%	105
		Online	99.19%	4.47%	61
	Undergraduate	Face-to-Face	98.98%	4.55%	952
		Hybrid	99.44%	1.88%	24
		Online	99.76%	1.38%	66

Table A.1

(continued)

SFF Question	Course Level	Instructional Method	Mean	Std. Deviation	*N*
Q11 Response Rate	Graduate	Face-to-Face	99.75%	1.51%	403
		Hybrid	99.92%	0.81%	105
		Online	100.00%	0.00%	61
	Undergraduate	Face-to-Face	99.41%	5.07%	952
		Hybrid	99.68%	1.57%	24
		Online	99.64%	1.82%	66

Table A.2

Summary of ANOVA for SFF Questions Response Rate

Source	Mean Square	F	Sig.	Partial Eta Squared
Between Subjects Effects				
Instructional Method	3137.03	60.57	.000	.070
Course Level	70.75	1.37	.243	.001
Instructional Method * Course Level	17.62	.340	.712	.000
Error Between	51.79			
Within Subjects Greenhouse-Geisser				
Source	Mean Square	F	Sig.	Partial Eta Squared
SFF Questions	1518.41	69.15	.000	.041
SFF Questions * Instructional Method	1258.03	57.29	.000	.067
SFF Questions * Course Level	41.16	2.75	.002	.002
SFF Questions * Course Level * Instructional Method	22.89	1.04	.353	.001
Error Within	21.96			
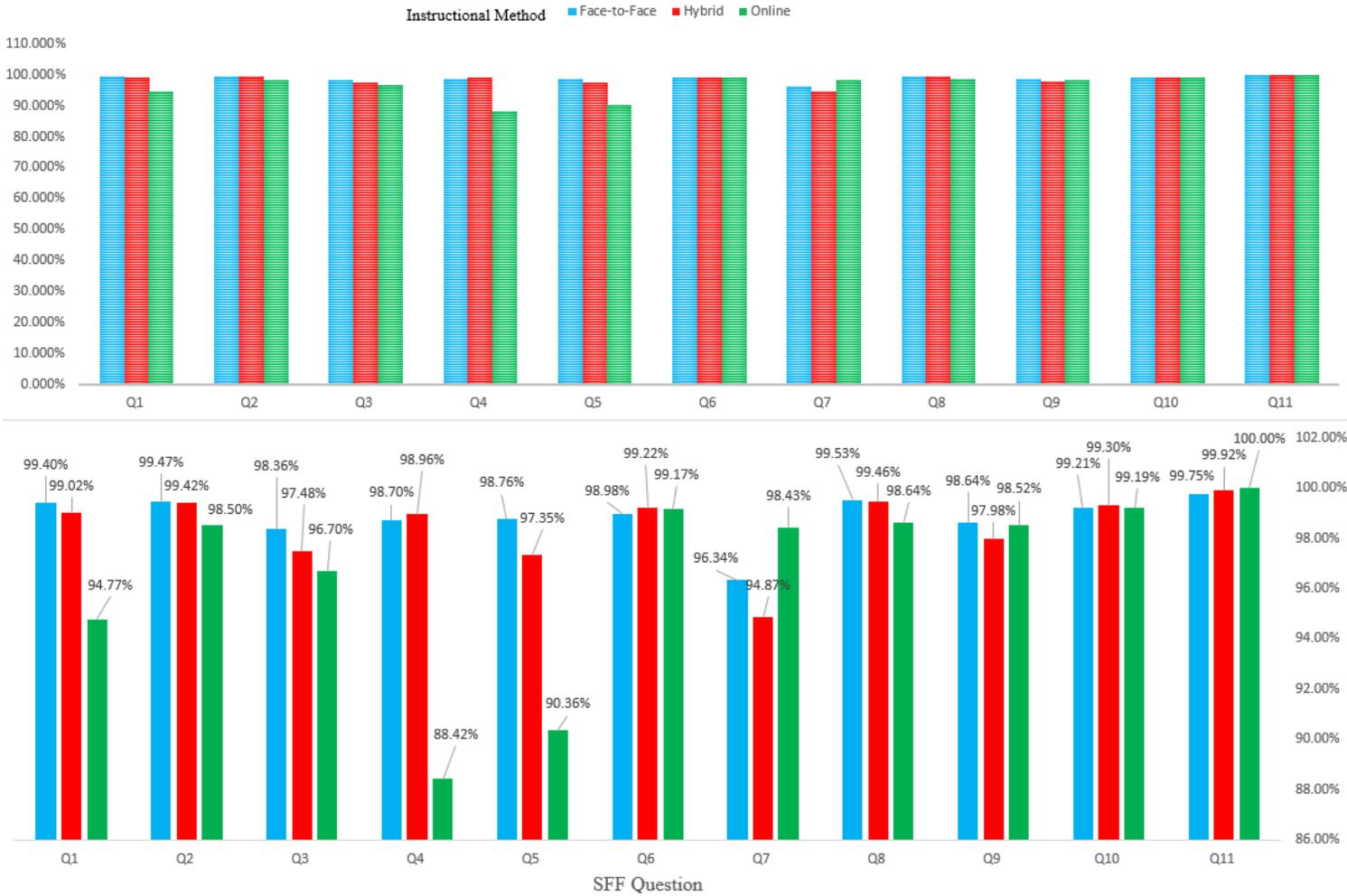
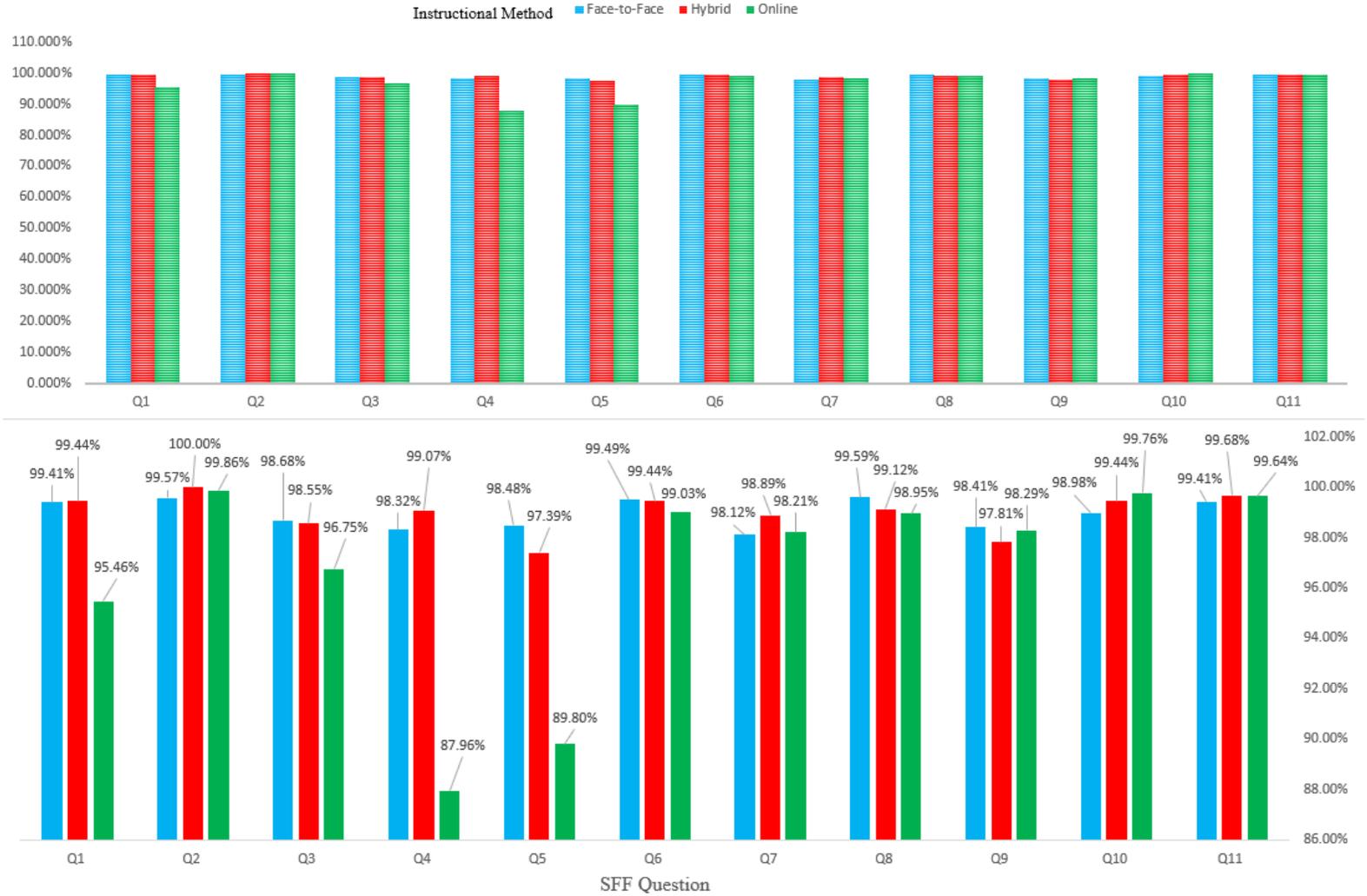Figure A.1 Graph of Means of SFF Question Response Rates for Graduate Courses

Figure A.2 Graph of Means of SFF Question Response Rates for Undergraduate Courses

Table A.3

SFF Questions Response Rate Means and Standard Deviations

Dependent Variable	df	Mean Square	F	Sig.	Partial Eta Squared
Q1 Response Rate	2	1060.896	92.694	0	0.103
Q2 Response Rate	2	6.486	0.886	0.413	0.001
Q3 Response Rate	2	232.731	12.430	0	0.015
Q4 Response Rate	2	6222.229	229.895	0	0.222
Q5 Response Rate	2	4207.723	148.738	0	0.156
Q6 Response Rate	2	3.574	0.443	0.642	0.001
Q7 Response Rate	2	275.959	6.832	0.001	0.008
Q8 Response Rate	2	35.358	6.050	0.002	0.007
Q9 Response Rate	2	16.849	0.711	0.491	0.001
Q10 Response Rate	2	14.962	0.982	0.375	0.001
Q11 Response Rate	2	11.865	0.740	0.477	0.001

Overall Wilks' Lambda = .681, p = .000. Partial eta squared = .175.

Table A.4

SFF Questions Response Rate by Instructional Method Means and Standard Deviations

SFF Question	Instructional Method	Mean	Std. Deviation	*N*
Q1 Response Rate	Face-to-Face	99.40%	2.73%	1355
	Hybrid	99.10%	3.00%	129
	Online	95.13%	7.54%	127
Q3 Response Rate	Face-to-Face	98.58%	3.82%	1355
	Hybrid	97.68%	6.26%	129
	Online	96.73%	6.50%	127
Q4 Response Rate	Face-to-Face	98.43%	4.11%	1355
	Hybrid	98.98%	3.36%	129
	Online	88.18%	12.35%	127
Q5 Response Rate	Face-to-Face	98.57%	3.97%	1355
	Hybrid	97.35%	5.67%	129
	Online	90.07%	12.62%	127
Q7 Response Rate	Face-to-Face	97.59%	6.27%	1355
	Hybrid	95.62%	8.60%	129
	Online	98.32%	4.23%	127
Q8 Response Rate	Face-to-Face	99.57%	2.21%	1355
	Hybrid	99.40%	2.30%	129
	Online	98.80%	4.08%	127

rates between face-to-face and hybrid instructional methods, there was were no significant differences found between these groups for the other questions. A significant lower response rate was found for all of these quesitons when comparing face-to-face and online learning as well as hybrid and online learning.

These results may best be explained by reviewing each of the questions. For instance, question one – "I came well prepared for class" – and question three – "The instructor was well organized and prepared for class" – may be considered invalid for online learning students where it may be unclear when a specific class meeting begins or ends. Hybrid students, who participate in some scheduled meetings, may be more likely

to respond as a reflection of those scheduled "class" gatherings. Similar issues may explain the lower response reates for question four – "The instructor was conscientious in meeting class and office hour responsibilities" – and question five – "The instructor promoted a classroom atmosphere in which I felt free to ask questions" – both of which appear to ask students to reflect on experiences related to specific times and/or physical meetings or spaces.

The differences in the response rates for question seven – "So far the instructor has applied grading policies fairly" – may be of more interest, considering that hybrid was lower than both face-to-face and online, which had the highest response rate. The fact that these differences are no longer significant when grouped by course level may indicate that one or the other involves assignments that are graded later in the term, after the majority of SFFs are completed. If students have not yet received grades on assignments that make-up a significant portion of their overall grade may result in skipping this question.

The differences in the response rates for question eight – "The instructor taught this course well" – is perhaps the most interesting. The difference was most notable for graduate level courses with online learning lower than either hybrid or face-to-face. This may reflect some level dissatisfaction with asynchronouse online courses that do not include any scheduled virtual or in-person meetings. Or perhaps students may not have felt they could evaluate the performance of an instructor with whom they did not engage with for direct instruction. Students who received all instruction mediated via an instructional platform, such as Canvas, may not have felt the presence of the "instructor" as actively engaged in "teaching" and therefore chose not to evaluate this experience. Further research may provide additional insight into this relationship. However, this difference in response rate did not seem otherwise relevant to the overall results as presented in this study.