3-2023

Student Performance in Traditional In-Person Vs. Online Sections of an Introductory Graduate Mathematics Course

Lauran E. Kittle

STUDENT PERFORMANCE IN TRADITIONAL IN-PERSON VS. ONLINE SECTIONS OF AN INTRODUCTORY GRADUATE MATHEMATICS COURSE

THESIS

Lauran E. Kittle, Captain, USAF

AFIT-ENC-MS-23-M-003

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENC-MS-23-M-003

STUDENT PERFORMANCE IN TRADITIONAL IN-PERSON VS. ONLINE SECTIONS OF AN INTRODUCTORY GRADUATE MATHEMATICS COURSE

THESIS

Presented to the Faculty

Department of Mathematics and Statistics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Mathematics

Lauran E. Kittle, BA

Captain, USAF

March 2023

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT-ENC-MS-23-M-003

STUDENT PERFORMANCE IN TRADITIONAL IN-PERSON VS. ONLINE SECTIONS OF AN INTRODUCTORY GRADUATE MATHEMATICS COURSE

Lauran E. Kittle, BA
Captain, USAF

Committee Membership:

Maj Victoria R. C. Sieck, PhD
Chair

Dr. Christine M. Schubert Kabban, PhD
Member

Dr. Michael E. Miller, PhD
Member

AFIT-ENC-MS-23-M-003

Abstract

The growth of technology impacts nearly every aspect of everyday life, to include education and learning. The availability of distance learning (online) classes has increased drastically in the last few decades, expanding access to education for millions of people. However, it is imperative to consider exactly how the growth of technology impacts education – whether it is a positive, negative, or neutral impact. Previous research comparing distance learning and in-residence (traditional) classes have widely mixed, disparate conclusions. This type of research, two-stage analysis, and modeling has yet to be conducted on a graduate school level. For this reason, a detailed look into how student performance is affected by different classroom environments at the Air Force Institute of Technology Graduate School of Engineering and Management (AFIT-EN) is conducted. This study analyzed two data sets: the first included observational data with multiple demographic variables. This data was then used to model student performance as determined by the student's final course grade. The second data set is experimental in nature as the class was meticulously designed to eliminate as many sources of bias as possible, with the goal of making the in-residence and distance learning sections of the course mirror each other. The results of both analyses indicated that when the sources of bias are controlled, either through variables or design of experiments, learning environment does not impact student performance. These results can be applied to other courses and schools, while also assisting professors in designing their courses and students in choosing which course section is right for them.

Acknowledgements

I would like to recognize the invaluable assistance of my faculty advisor, Major Victoria Sieck. She listened to the topics I was interested in, established a research topic, and then found courses and professors who were willing to participate. She gave this study the focus it so desperately needed in my own hands. She took the time out of her busy schedule to walk me through statistical concepts and processes that were new to me. But not only did Major Sieck provide the academic and educational support necessary for any research, she gave me the confidence and organization needed to finish my research – through our meetings and through watching her juggle her own home and work life.

I would also like to thank my other committee members, Dr. Schubert Kabban and Dr. Miller. Through the multiple classes I took with both of them and through their invaluable insights and suggestions for this research, they created a foundation of academic support required for success.

Finally, this research would not be remotely possible without the support and love of my family. My parents have always encouraged me in my dreams and decisions, even when they knew it meant they wouldn't see me for months at a time. To my husband: since the moment I met you, you have been my biggest supporter. Like many before us, this adventure has not been an easy road, but I know as long as you are by my side, I am capable of anything I set my mind to. Through your endless patience and unconditional love, you give me and our family everything we need to grow. To my son: you have taught me more than I will ever teach you and I am beyond grateful to have you in my life. I can't wait to see you grow into the amazing human I know you will be.

Lauran E. Kittle

Table of Contents

Page

List of Figures

Page

List of Tables

STUDENT PERFORMANCE IN TRADITIONAL IN-PERSON VS. ONLINE SECTIONS OF AN INTRODUCTORY GRADUATE MATHEMATICS COURSE

I. Introduction

1.1 Chapter Overview

This chapter explores the effects technology has on education and student performance, particularly at post-graduate schools such as the Air Force Institute of Technology (AFIT), a well-known research institution. The general issue is presented, which serves as the driving factor for a set of research questions. An overview of the study that aims to answer those questions follows.

1.2 General Issue

The growth rate of technology is consistently increasing: the accessibility, availability, processing power, and costs of technology are wildly different now than 20, or even 2 years ago. In many application areas, such as the medical field, these technology advancements are clearly a positive factor (Cristea, 2020; Kruse & Beane, 2018). But this is not the case in all fields; in some, the application of technology advancements is not so plainly beneficial. One such field is that of education and learning.

The concept of applying technology in the classroom seems promising. It can save time, equip students with necessary digital skills for the future, and be personalized to fit individual students and teachers. But does it work in practice? Has education changed to embrace technology advancements? Instructors lecturing at a podium has been the default in education for centuries, as shown in Figure 1, a painting from the Italian artist Laurentius de Voltolina from the 14th century illustrates the longevity of group lecture. As is common in today's classrooms, many of the students appear to paying attention and are actively taking notes, but others appear to be talking to their peers and one even appears to be sleeping.

Figure 1: 14th Century Painting by Laurentius de Voltolina

Yet while the concept of lecturing in an educational setting is still in use, there are many ways in which the field of education has advanced. Technology, including the advancements of public transport and motorized vehicles, expanded access to education for millions of people. Innovations such as the printing press made books cheaper and more common, allowing individuals other than nobility and elite societal members access to reading materials.

In more recent years, the response to the Coronavirus infection (COVID-19), forced an increased application of technology into the classroom. While schools all over the world shut down in an attempt to combat infection rates, many schools also sought ways to continue instruction without being in close contact or in-person. In the United States, more than 1,300 colleges and universities cancelled in-person classes or shifted to online-only instruction during the spring semester of 2020. By the fall semester of 2020, schools developed plans to merge in-person instruction with social distancing and online learning (Smalley, 2021).

Some students lacked the tools and resources for online learning and many colleges tried to assist in various ways. Some colleges attempted to combat the concerns of student performance in online instruction by opening libraries to provide reliable technology and internet connections to students. Others distributed devices and mobile hotspots to students (Smalley, 2021).

To support the transition to online learning, instructors also had to shift their primary methods of teaching. Online video chat services such as Zoom, Google Hangouts, or Microsoft's Skype were popular choices due to their functionality and availability to users. Some instructors arrived early and stayed late when hosting a class session to meet their students and reproduce opportunities for informal chats or questions. One-on-one meetings, either in person or online, were another way instructors adjusted to build relationships with students. During this time, online quizzes and surveys engaged students, online document sharing helped students complete papers, theses, and dissertations, and recorded videos were used for sharing lectures and student presentations. While most of these techniques, methods, and services were not new, they required many instructors to adjust to meet requirements (Cheong and others, 2021; D'Agostino, 2022; de Vries, 2021; Wiyono and others, 2021).

Additional contingency plans were made by a number of schools when faced with requirements to provide online or blended classes. For example, some schools shifted to a pass/fail grading schema in lieu of the standard A through F letter grades, which was intended to curb the stress students experienced in these unprecedented times, when many of them had never taken an online class (Smalley, 2021; Svrluga, 2020). Indeed, even the Graduate Record Examination (GRE), a standardized test designed to measure academic readiness for graduate school, became available for people to complete at-home. The Educational Testing Service (ETS), who creates and administers the GRE, introduced the at-home version in response to the pandemic and social distancing requirements. The at-home GRE is now offered everywhere the test is normally available - as long as the individual has met the equipment and testing room requirements (Kowarski, 2021; *Taking the GRE General Test at Home*, n.d.). However, some schools and departments decided to suspend the requirement for GRE scores, either temporarily or permanently, fueled by the potentially burdensome requirements for at-home test takers and concerns that the GRE doesn't predict student success in graduate school (Hu, 2020). AFIT was one such school that waived the requirement for GRE scores for admission during the first peak of COVID-19 in 2020 (AFIT ENE Admissions, personal communication, October 31, 2022).

These technological changes had varying degrees of success for each institution and each student. Some students excelled in their new online or blended environment, while others struggled and fell behind. These various levels of success are further discussed in Chapter II, and prompt the question: is there a balance when blending technology and education so that the benefits to students are maximized?

1.3 Research Objectives

This research seeks to determine the impact instructional technology has on student performance, as determined by final grades, at the graduate school level (specifically at AFIT-EN) and to use that data to understand the balance of technology for individual teachers and students. Time (as technology grows), student demographics (such as age and background information), and instructor changes should all be considered as potential factors. In particular, any influence the unique nature of the Air Force Institute of Technology Graduate School of Engineering and Management may have on combining technology and education will be assessed.

To satisfy the objectives of this research, the following questions are presented, investigated, and answered:

1. How is student performance affected by traditional in-person instruction versus online instruction?
2. How does controlling for other variables and biases affect the impact instructional technology and learning environment have on student performance?
3. How does AFIT differ from other graduate schools, and how do those differences affect the first two questions?

Two data sets will be used to answer these questions, both of which come from a four-credit, graduate-level, introductory mathematics course. These two data sets are introduced further in the following section, which also details how they will be analyzed. Both the first and second data sets used in this research come from AFIT's Graduate School of Engineering and Management (AFIT-EN). The unique makeup of AFIT-EN, which is discussed more in Chapter II, is assumed to have an impact on student performance, before, during, and after the time spent at AFIT.

1.4 Research Focus

As previously stated, the focus of this research is on the impact technology use has on student performance, as determined by final grades, both continuous scores and letter grades. Examination grades are provided and analyzed for trends and inferences as well. Course information (syllabus summary, format, quarter, and instructor) and student demographics (age, Department of Defense (DoD) affiliation, degree, and department) are further included to determine correlations and potential causations with student performance.

To conduct these analyses aimed to answer the research questions, this research utilizes various mean, median, variance, and distribution comparison tests between online and in-person final grades, as well as extensive model building to predict students' final continuous grades. The first data set is comprised of nine previous quarters of the given mathematics course. The second data set, from the most recent quarter of this course, is presented and comparison tests, both within this data set between online and in-person final grades and also compared to the first data set, are completed. Lastly, model building on the second data set is conducted and compared to the model from the first data set. A combination of JMP Pro 15, Microsoft Excel 2019, and R 4.2.2 are utilized for all analysis in this research.

This type of research and analysis is unique from existing literature in a few ways. First, this extensive analysis on student performance in regards to instructional technology has not been done at AFIT-EN. Since it is assumed that the unique makeup of the graduate school will have an impact on student performance, this research is imperative in discovering what these impacts are and why these impacts occur. This will pave the way for future research at AFIT-EN, which may in turn allow for similar studies and comparisons within similar institutions.

Additionally, the two-part analysis used in this study, in which an observational study is completed on decades of data and influences the design of a follow-on experimental study, has so far not been considered in the literature to the best of our knowledge. Chapter II reviews a myriad of observational studies with various conclusions, and even a few experimental studies. However, none of them have the combination of the observational research and the application of the results of that research as an experimental study, as seen in this study.

Lastly, the type of modeling used in this study is rarely seen. The analysis of student performance is often completed, but creating a prediction model with numerous course and student demographic variables is uncommon, although increasing in prevalence. Yet, understanding how students may perform, given specific characteristics, is essential for instructors to guide their students' success, regardless of the learning environment.

1.5 Investigative Review Board

Due to this research involving data collected about individuals and interviewing professors, a request to exempt this research from institutional review board oversight of human subjects' research was sent to the Air Force Research Laboratory Institutional Review Board (IRB). The IRB determined on 3 August and 23 September 2022, respectively for the first and second studies and corresponding data sets, the research involved an established accepted educational setting with normal educational practices that would not negatively impact students or instructors. Furthermore, it was determined the demographic information that was collected was not beyond the limits of normally collected classroom data.

Subsequently, the AFIT Office of Institutional Research collected the student demographic data and deidentified all data to ensure the privacy and anonymity of students and instructors to comply with AFIT policy that follows federal laws, such as the Family Educational Rights and Privacy Act (FERPA). This data was then provided to the author for statistical review and analysis.

1.6 Preview

This document is systematized in the following manner: Chapter II gives a detailed look into literature and studies with similar research questions and backgrounds and discusses the methods utilized in this research. Chapter III provides AFIT-EN information, summarizes the instructor interviews conducted, and presents the analysis on the first data set (covering 9 previous quarters). Chapter IV details the experimental study conducted with a full analysis of the second data set collected and how it compares to the first data set. Finally, Chapter V reviews conclusions from the study and presents recommendations.

II. Literature Review and Methodology

2.1 Chapter Overview

This chapter analyzes previous research and findings on various instructional technology and online training methodologies. This review includes descriptions, summaries, and an in-depth study of previous research conducted in the field of instructional technology. Online learning and records of previous studies are also detailed to include assorted systems and methods, education levels, and subjects. This is followed by a detailed look at previous studies, directly comparing traditional teaching methods to methods that employ modern technology. This chapter then concludes with a comprehensive summary of the methods and procedures utilized throughout the remaining study.

2.2 Instructional Technology

Instructional technology (IT) refers to the use of various computer hardware, systems software, and applications software to facilitate student learning. Ideally, these technologies are designed to create engaging and effective learning experiences to improve student performance. A more formal definition of instructional technology is: "the theory and practice of design, development, utilization, management and evaluation of processes and resources for learning" (Seels & Richey, 2012). While this definition seems incomplete as it doesn't specifically mention technology, it is still widely accepted as the field definition since its introduction in 1994 (*FAQ: Instructional Technology Program, Watson College of Education at the University of North Carolina Wilmington*, n.d.; Li and others, 2009). Even with newer definitions constantly being introduced, such as one from Lane and others that details instructional technology as: "the design, development, use, management, and evaluation of the process of learning mediated by technology applications" (2019), many experts and authors prefer the 1994 definition. Gagne describes this divide within the field by saying the definitions of instructional technology fall under one of two types: one that equates it with a particular set of instructional media (or audiovisual devices) and the other that equates it as a process (or systems approach) (1987).

Depending on how one defines instructional technology, it may also be referred to as educational technology, although some define educational technology as the procedure or theory of using the instructional technology tools (*Difference Between Educational and Instructional Technology*, n.d.; Seels & Richey, 2012). Indeed, the Association for Educational Communications & Technology (AECT) issued the following definition for educational technology: "the study and ethical application of theory, research, and best practices to advance knowledge as well as mediate and improve learning and performance through the strategic design, management, and implementation of learning and instructional processes and resources" (*Association for Educational Communications and Technology*, n.d.). This definition, introduced in 2008, is cited by several experts and is described as building on the definition proposed for instructional technology in 1994 (Ahmadigol, 2016; Januszewski, 2008). For the purposes of this research, instructional technology and educational technology will be used interchangeably, most closely related to the definition by Lane and others (2019), but it is important to note that technological advances and societal changes will keep these definitions dynamic.

Types of instructional technology can fall into the following five categories: synchronous, asynchronous, blended learning, linear learning, and collaborative learning

(*5 Different Types of Educational Technology | TL Dev Tech*, 2022). These types are not mutually exclusive, nor must they all be employed together. For example, some authors may combine synchronous and asynchronous or omit blended learning ("What Are the Different Types Of Educational Technology," 2021). For this research, all five categories of instructional technology will be considered and examples of each are provided for further detail.

Synchronous is defined as occurring at the same time; when applied to learning, this means all students are participating in the lesson and connecting in real-time. Synchronous learning can be in person or online, and may also be referred to as live or real-time instruction (Chen and others, 2005). Examples of instructional technology tools that can be applied to synchronous learning are video conferencing, chat rooms, and interactive whiteboards (SMART Boards) (Park & Bonk, 2007).

Asynchronous is the opposite: not occurring at the same time. This type of learning enables students to access class material and complete coursework at any time, at any place, and at their own pace (Jaffee, 1997). Asynchronous learning enables students to connect with each other around the globe and is beneficial for students with demanding schedules. Examples of instructional technology here may include blogs, emails, online textbooks, and pre-recorded audio or video lectures (Nissen and others, 2014). Chat rooms, while previously mentioned as synchronous, are a type of instructional technology that can be considered as either synchronous or asynchronous, depending on how it is implemented by the teacher and utilized by the students. For example, a chat room that is used in conjunction with a live lecture would be considered synchronous, but an online discussion board where students interact with each other over the case of many days or weeks would be considered asynchronous.

The third category is blended learning, where students and teachers meet in-person and also have online portions. This type of instructional technology can be considered as a combination of synchronous and asynchronous learning. The terms blended and hybrid learning are used interchangeably for the purposes of this study and are further detailed in the following section titled "Flipped and Hybrid Classrooms." Examples here include any tools from both synchronous and asynchronous learning, such as face-to-face teaching and computer technology (Hockly, 2018; Sharma, 2010). However, a significant difficulty with blended learning is that it requires instructors to be well-versed in both synchronous and asynchronous methods of information delivery.

The fourth category of instructional technology is linear learning. In a traditional in-person classroom without the use of technology, linear learning takes the form of lesson plans that consecutively follow a textbook's chapters, without deviations or additions in content structure. Linear learning in educational technology focuses on Computer-Based Training (CBT) and requires students to complete their training sequentially. Here, information is pushed to student devices, generally in the form of videos, animations, and readings ("What Are the Different Types Of Educational Technology," 2021). Bedwell and Salas aptly summarize CBT as a self-contained and interactive program designed for self-paced instruction that uses student-controlled features, predesigned material, required responses, and feedback (Bates, 2005; Bedwell & Salas, 2010; Rosenberg, 2005). Many linear learning programs allow instructors to track student progress and efficiently identify areas which may need further support (Bedwell & Salas, 2010; Lee and others, 2016).

Lastly, collaborative learning entails a group of students working together to complete coursework: solving a problem, completing a task, or creating a product (Laal & Ghodsi, 2011; Laal & Laal, 2011). Collaborative learning can be done in-person or online; the latter relies heavily on discussion boards, chat rooms, collaborative software (e.g., Google Docs), and video conferencing. This type of online learning may also be referred to as Computer-Supported Collaborative Learning (CSCL) (Lee and others, 2016; Strijbos, 2011).

In conclusion, within the overview of these five categories, examples are provided of instructional technology and how they may be used within the category. While not exhaustive, each of the examples is representative of their potential uses in the classroom setting. A deep dive of the third category is now conducted to further understand how instructional technology is employed to support blended learning.

2.3 Flipped and Hybrid Classrooms

As previously identified, this section details the third category of instructional technology, blended or hybrid learning, and compares it to flipped learning. Flipped and hybrid classrooms have become more popular recently and are looked at separately, although they can overlap. Both have broad descriptions, causing their interpretations and applications to vary widely.

Hybrid classrooms have a very nonspecific definition that describes an environment in which students are face-to-face for parts of the class and are online for others (Ward, 2004). This online portion can include virtual class participation or another form of technology such as online tests. The concept of hybrid classrooms is well established and they are often tailor-made for each course and set of students. Hybrid classrooms may also be referred to as blended classrooms, integrated learning, or multi-method learning.

Flipped classrooms entail students completing readings or watching lectures before scheduled class time. Then during class, students have the support of the instructor to complete the more difficult concepts and applications, and are able to ask questions (Brame, 2013). Flipped classrooms do not require the use of instructional technology, but when that technology is used for the work before each class session (as in the case of watching lectures), it would also be classified as a blended or hybrid classroom. In essence, lower levels of cognitive work are done outside of class with or without technology and the higher forms of cognitive work are done in class with instructors and peers (Gundlach and others, 2015). This type of learning prevents cognitive overload by breaking down lectures into more manageable pieces. It also capitalizes on the time spent during the in-person classes (Karaca & Ocak, 2017). Flipped classrooms may also be referred to as inverted classrooms or peer instruction, all of which imply a pedagogy-first approach to teaching.

A deeper look at how and when flipped classrooms and hybrid classrooms intersect one another is now conducted. When flipped classrooms require instructional technology or online participation before class (i.e., in the form of various online readings, lectures, or assignments), they may also be defined as a hybrid classroom for their use of technology. Figure 2 was created to provide a visual of this scenario as a Venn diagram.

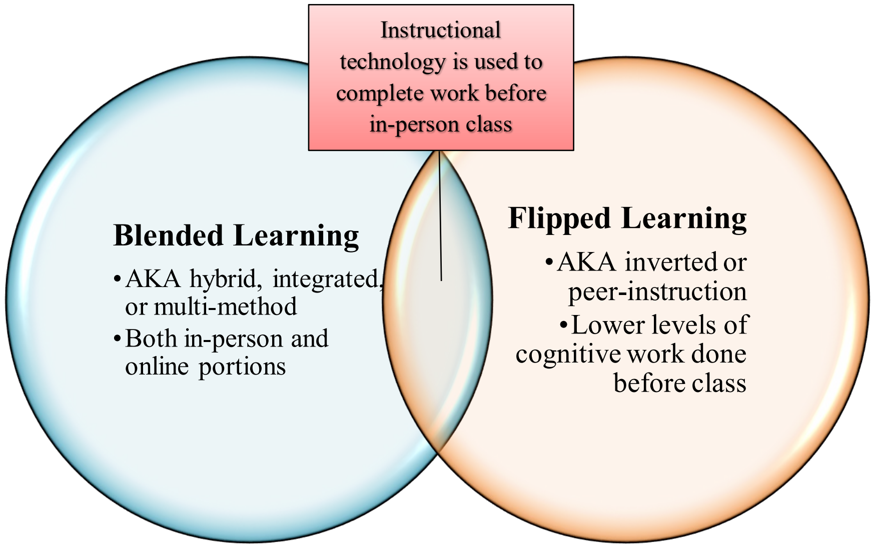


Figure 2: Venn Diagram of Blended vs Flipped Learning

For example, the United States Air Force has previously used flipped and hybrid learning for Chemical, Biological, Radiological, and Nuclear (CBRN) mandatory training. Individuals were required to complete the CBRN Defense Awareness CBT before attending an in-person training which ensured they were properly oriented with their gear, knew the proper procedures pre- and post-attack, and put their learned skills to the test in a realistic environment (Morales, 2017).

Conversely, there are situations in which flipped and hybrid classrooms do not overlap. For example, a classroom set up in a flipped manner in which the lower level of cognitive work completed before class includes reading a textbook with no other instructional technology, would not be considered a hybrid classroom. Similarly, a blended classroom in which the technology is used throughout the course, such as online homework or online exams, would not be considered a flipped classroom.

As will be shown in a review of the literature, these types of classrooms (with or without use of technology) can increase student performance; therefore, this will be considered further in Chapter IV as a key element of the experimental study. The following sections describe the benefits and limitations of instructional technology use in any of the aforementioned categories.

2.4 Benefits and Advantages

The application of instructional technology has seemingly limitless potential. In today's world, technological advances can be exciting and engaging to the vast population. Studies have shown that, in general, as students' engagement increases, motivation and academic performance increase as well (Dogan, 2015; Sun, 2014). For example, Virtual Reality (VR) immerses the user in such a way as to give the perception of being physically present in the non-physical world (Lum and others, 2020). VR can also provide experiences that traditionally require extensive resources or difficult, if not impossible, to acquire materials, such as history exposures and science experiments, benefitting students that may not have the opportunity to acquire those experiences.

Instructional technology also enables students and instructors to collaborate and share resources, both with each other and with their peers. Recorded lectures can be paused and rewatched to support further instruction or practice. Software such as Moodle, Socrative, Google Class, and Microsoft Teams allows students to interact directly with instructors and peers, without the concern for time, distance, or space availability that traditional classrooms and office hours have (Limantara and others, 2019). Furthermore, instructors can both provide and receive timely and pertinent feedback with the use of instructional technology. For example, online surveys and polls, clickers and questionnaires deliver quick responses which instructors can leverage to improve current students' experiences and restructure possible future classes. As an added benefit, online feedback systems are reliable and can feel low-threat to students, encouraging greater participation among students of widely varied backgrounds and abilities (Hatziapostolou & Paraskakis, 2010).

Lastly, many scholars believe that individuals have a myriad of different learning styles. A number of students learn best through watching and listening, some by reading and writing, and others by doing and moving (Zapalska & Brozik, 2007). Some studies have shown that in using multiple teaching approaches that cater to different learning styles (known as employing a multimedia effect), students' academic success can increase (Ilhan & Oruç, 2016). In understanding students' learning styles, instructors can tailor their classes and teaching methods to better suit their students. Content can be presented in lectures synchronized with PowerPoint presentations to fit the needs of auditory and visual learners. Discussion groups and supplemental materials from external websites can be provided for individuals who prefer to read and write. Course projects, virtual field trips, and videos may help those who utilize all their senses and prefer "hands-on" experiences. Ultimately, the extensive options and inherent possibilities of instructional technology and online learning enable instructors to customize courses to accommodate individual learning styles (Viorica-Torii & Carmen, 2013).

Ultimately, how and why technology is employed must be considered to obtain the aforementioned benefits and advantages. Hernández González & Blackford state that proper application of classroom technology, along with appropriate interaction techniques for each course, can have a positive impact on student engagement. This in turn, improves intrinsic motivation, which has been shown to increase student performance and scores (2022).

2.5 Limitations and Disadvantages

Conversely, the limitations and disadvantages of instructional technology must be considered as well. At the forefront of many opponents' arguments is the potential for distraction caused by utilizing technology in classrooms. Because of the engaging aspect of technology, students may be apt to utilize technology for entertainment during class time, disassociating their attention (Limantara and others, 2019; Pappano, 2012; Pienta, 2013). This can be true for fully online courses as well; as students attempt to set aside time to review class materials, they may find themselves utilizing their devices for leisure instead.

Additionally, the cost of instructional technology may not offset the potential benefits. Initial and ongoing acquisition of proper bandwidth amounts, hardware and devices, software, and maintenance requirements can be overwhelming for students, instructors, and institutions. Obtaining the funding for instructional technology provides yet another challenge for institutions (Groff & Mouza, 2008).

The overall difficulties in using these instructional technologies must be considered as well. There is a significant learning curve which must be mastered to properly apply and instruct with instructional technology, requiring teachers to be adequately prepared. Although results depend on the specific teacher preparation programs, research has shown that instructor preparation can positively impact student achievement (Boyd and others, 2009; Viorica-Torii & Carmen, 2013). However, many teachers report they do not feel prepared to instruct with technology upon completing their teacher education programs (Korucu-Kis & Ozmen, n.d.). Technology failures further strengthen the argument against using technology in classroom settings. Both student and teachers alike need to have contingency plans if their Internet, devices, or software are inoperable (Reid & Lambert, 2014). Yet even with contingency plans, there may be no choice but to cancel or reschedule a class, redo an assignment, or even submit incomplete or missing work as a result of data loss.

Finally, it can be difficult to determine facts from fiction with the misinformation that can be found online. Del Vicario and others stated: "the wide availability of user-provided content in online social media facilitates the aggregation of people around common interests, worldviews, and narratives" (2016). But misinformation is not found only on social media; the Internet is available to the vast majority of individuals, and anyone has the opportunity to create a website or blog to publish their own ideas and opinions. While there are ways to combat this issue, it is nonetheless another limitation to incorporating instructional technology into the classroom.

2.6 Why Distance Learning?

When considering the advantages and disadvantages of instructional technology, one can also examine why students decide to engage in instructional technology and take online classes when the course is offered both online and in-person. By doing so, professors and academic advisors can potentially help guide students to which class may be best suited for them.

McPartlan and others summarized the findings of several studies as to who is more likely to choose online courses: “Women are generally more likely to enroll in online courses than men. Additionally, online students are more likely to be older, employed, and single parents, reflecting the flexibility desires for completing studies alongside employment and family responsibilities” (2021). This gives an idea of the type of individual most likely to enroll online, but a closer look at the reasoning is imperative.

To understand the reasoning, student motivation must be considered. However, it is rarely, if ever, accounted for as a pre-existing condition in studies comparing student performance. This may be by virtue of the difficult-to-capture nature of an individual’s motivation. Surveys pre- and post-course are a common way to capture motivation. In some studies, the flexibility of online classes is a recurring theme; the ability to complete coursework at personal convenience is no minor matter. Other frequently repeated motives include: long commute to school, no available spots in the face-to-face course, self-regulation to move at a slower/faster pace, and decreased course costs (McPartlan and others, 2021; Vanslambrouck and others, 2018).

Understanding students’ motivation, and therefore their personal situations and perceptions, enables instructors to anticipate their needs. For example, a single employed mother may need more flexibility if classes meet synchronously. Children get sick often and if she has no other support systems, she may turn in assignments late. Being considerate of an individual’s unique situation can mean the difference between success and failure, finishing a course or having to drop out of it. Next, how students perform in different classroom environments is discussed, as well as which factors mentioned in this section may influence performance, such as age, gender, and student hours/course workload.

2.7 Comparing Student Performance Among Various Classroom Environments

While there is a myriad of research completed on comparing test scores of students in various course environments, this section will focus mainly on those studies that dealt with undergraduate and graduate mathematics and statistics courses, as this is the focus of this research. However, there are other science, technology, and engineering courses included to provide a wider breadth of studies and to showcase how instructional technology may compare in relation to different courses. These studies can broadly be categorized based on their results into three types: studies that found no statistical difference between course environments, studies that found instructional technology courses performed better, and studies that found instructional technology courses perform worse. Each of these results will be considered separately below.

2.7.1 No Significant Statistical Difference Between Course Environments

The studies in this section concluded there was no significant statistical difference in student performance between course environments that did and did not involve instructional technology usage. Unless otherwise stated, these studies were retrospective, using existing information and data that was previously collected for other purposes.

While the research is now dated by two decades, Merisotis and Phipps summarized the results of about 40 studies and concluded: "…regardless of the technology used, distance-learning courses compare favorably with classroom-based instruction and enjoy high student satisfaction" (1999). The substantial changes in technology, and specifically instructional technology, over time were discussed in Chapter 1, but this type of meta-analysis is important to consider on a large-scale approach.

In a more recent study from 2021 by Hoffman & Elmi, a strict comparison between asynchronous online and traditional in-person learning environments was conducted on a graduate-level introductory biostatistics course from a private university in Washington, D.C. Although this study used existing data, several possibly significant sources of bias were able to be controlled because of the unique nature of the study: the same instructor taught both sections, both sections utilized the same basic weekly structure, and student demographic characteristics and previous academic performance were taken into account to accomplish analyses of students with certain common characteristics. The average quiz and exam scores, which were the responses for the study, were slightly lower for online students than scores for students in the in-person section. However, the authors ultimately concluded the differences were not significant using averages, Likelihood Ratio tests, scatterplots, and confidence interval comparisons (Hoffman & Elmi, 2021).

Another study with some similar conditions (same instructor and same course) came to a comparable conclusion for a graduate applied statistics course. Stephenson (2001) took the grade points averages (GPA) for in-person and online students for 10 semesters and compared the averages. He then used a Chi-squared (χ^2) test of independence to determine statistical difference between the letter grades and the learning environment of the students, without consideration of demographic characteristics. Ultimately, it was determined that the letter grades were independent of whether the student was enrolled in the in-person or online section (Stephenson, 2001).

Another study conducted by Gundlach and others in 2015 compared three different combinations of learning environments, with similar results. In this study, web-augmented traditional lecture, fully online, and flipped sections of an undergraduate statistical literacy course were also found to perform statistically similar in homework and projects. However, the students in the web-augmented traditional section, which used a notable amount of web-based technology to complement the face-to-face portion, scored statistically significantly higher on average on all examinations. All the conclusions for homework, projects, and examinations were made utilizing Chi-squared tests, Analysis of Variance (ANOVA) comparisons, and paired-samples *t*-tests. Additionally, many students in the course stated they had never taken an online or flipped course, which the authors state may imply that being in an unfamiliar learning environment could have negatively impacted their scores (Gundlach and others, 2015).

All of the aforementioned studies had one specific limitation in common that may create bias in the analysis: there was no randomization of students in each section, meaning students could choose which section to enroll in. It is possible some students only chose the section that fit their schedules, not which one they felt they would best perform in. This may imply further unmeasurable influential variables that are associated with both learning environment and academic performance.

2.7.2 Instructional Technology Courses Perform Better

Conversely, the authors of a meta-analysis study analyzing the performance of students in traditional lecture and flipped sections using instructional technology of undergraduate introductory statistics courses found the students from the flipped sections that used instructional technology significantly outperformed those in the traditional lecture sections. Standardized mean differences were calculated to make those conclusions. Similar to the previous studies, the non-randomization of student placement was also a concern in this study, as was the potential for instructor preference for a particular format. Nonetheless, this study exhibits the potential flipped learning has for increasing student performance (Farmus and others, 2020). Similar studies have also shown the potential for increased student performance and long-term retention by “flipping the classroom” in mathematics and statistics courses (Strayer, 2012; Wilson, 2013; Winquist & Carlson, 2014). Other studies in physics courses have found a flipped classroom can produce learning gains of nearly two standard deviations or more compared to traditional learning (Deslauriers and others, 2011; Hake, 1998). All of these studies used various applications of technology for the lower level of cognitive work outside of the classroom, from quiz submissions to viewing pre-recorded lectures to online discussion boards.

Further studies specifically compared and analyzed online and in-person courses for college algebra and business statistics, likewise concluding online students performed better (Dutton & Dutton, 2005; Lazari, 2018). The business statistics study by Dutton & Dutton (2005) utilized the same instructor in the same semester, with nearly identical weekly academic structures. Demographic data such as GPA, course hours, and major was also included for analysis. The college algebra study by Lazari (2018) had the same instructor across multiple semesters with an expedited pace for the online sections (although the overall structure was the same). Both studies used a combination of descriptive statistics, regression analysis such as two-tailed t-tests, and standard hypothesis tests to determine significant differences and similarities between the two course environments, within undergraduate institutions.

2.7.3 Instructional Technology Courses Perform Worse

On the opposite side of the spectrum, a few studies have shown students enrolled in the in-person section have performed better than students enrolled in the online section. Tanyel and Griffin conducted a study in 2014 utilizing ten years of data from various undergraduate courses comparing in-person and online data. To be included, courses had to be taught by the same instructor, both online and face-to-face during the same semester. Course subjects included arts and sciences, business and economics, and education; the preponderance of which came from arts and sciences. Using Chi-squared tests, they concluded students in face-to-face courses scored significantly higher than those in online courses (Tanyel & Griffin, 2014).

A large-scale study from DeVry University's undergraduate program also concluded students in online courses do not perform as well as they would have in the same course, in-person. The data for this study included over 230,000 students in 750 different courses, over the span of four years. The authors stated: "For the average student, taking a course online, instead of in a traditional in-person classroom setting, reduces student learning, as measured by course grades, and lowers the probability of persistence in college" (Bettinger and others, 2015). Although this study provided substantial evidence that online courses decrease student grades, it is noted that they cannot provide a full welfare analysis. For example, online courses at DeVry University are accessible to a much wider range of individuals than would be able to take in-person classes, whether due to location, cost, or time availability. The authors also mention that technology is consistently advancing, and these results are not an end state; changes could be made to decrease the variability between learning environments (Bettinger and others, 2015).

To combat the most recent widespread hurdle for academic institutions during the COVID-19 pandemic, one study analyzed the effect online learning had at the United States Military Academy at West Point. This study is unique in that the students were randomly placed in either the in-person or online section of a Principles of Economics course. Five hundred fifty students were randomized across 12 instructors in 38 different class sections of this required course during the Fall 2020 semester. Once fixed effects based on instructors, class time, and demographics are added, online instruction was shown to significantly reduce students' final grade by approximately one half of a letter grade. This variation was more pronounced for academically at-risk students. Ultimately, the differences in performance question the effectiveness of online courses for all students (Kofoed and others, 2021).

2.7.4 Literature Comparison Summary

This collection of existing literature and studies compare traditional and technological teaching methods, with three main conclusions: studies that found there is no statistical difference between course environments, studies that found instructional technology courses performed better, and studies that found instructional technology courses perform worse. Each study in the previous section has attributes that are shared and attributes that are different from this study. These may include: the year the data was collected, the number of courses, professors, or semesters, the type of institution or the scope of the study. In particular, studies conducted on graduate schools seemed to show no difference in student performance based on course environment, while studies from undergraduate schools were mixed. These attributes and conclusions for each study are summarized in Appendix A.

The parameters and variables found in these studies provided a starting point for developing a model for understanding how instructional technology impacts student performance. Furthermore, some of those parameters and variables were considered in Chapters III and IV, and even impacted the design of the study found in Chapter IV. Now the statistical process and methods that are used in this study are detailed, which were heavily influenced by the studies described in this section.

2.8 Statistical Procedures

The following subsections set forth the statistical procedures, approaches, and aspects employed in this research, throughout Chapters III and IV. The differences between categories of studies and corresponding statistical approaches are discussed first, followed by statistical procedures and considerations to include: Type I error, nested and crossed designs, outliers, and assorted statistical tests. The studies cited in Section 2.7 are mentioned throughout to provide examples of applications in the study topic.

2.8.1 Observational vs Experimental Studies

There are two main categories of research studies: observational and experimental. To fully understand the extent of the results, findings, and conclusions in the study, it is important to understand how the study was designed. In observational studies, researchers observe the effect of a given factor or treatment without controlling the variables or environment and record those observations. Generally, observational studies are less expensive; however, they may take longer to complete the observation period and the evidence provided to show a cause-effect relationship within factors is weaker than experimental studies (Khanna, 2020). The following studies are some of those discussed in Section 2.7 which were observational studies: Bettinger and others, 2015; Hoffman & Elmi, 2021; Merisotis & Phipps, 1999; Tanyel & Griffin, 2014. All of these studies utilized data that had previously been collected, without controlling the environment.

Conversely, in experimental studies, researchers introduce a factor(s) or intervention(s) randomly to a group and study the effects. The manipulation of the environment is what makes the conclusion for a cause-effect relationship stronger than in an observational study. Other key points of experimental studies may include: closely monitored, high cost, and often smaller and shorter than observational studies (Khanna, 2020). The following studies discussed in Section 2.7 were experimental studies: *Kofoed* and others, 2021; Stephenson, 2001. These studies controlled the environment in which the study was conducted; both were able to choose how the classroom environments were similar or different and the study by Kofoed and others in 2021 was also able to randomize the students in the class sections. The data set used in Chapter III is classified as data from an observational study: the data had already been collected and there was no manipulation of the environment. Conversely, the elements of the second data set are considered experimental and will be discussed in Chapter IV.

2.8.2 Parametric vs Nonparametric Statistics

One way to differentiate statistical procedures is to classify them as parametric or nonparametric. Parametric statistics has a number of assumptions that must be met to properly utilize and apply the tests to make conclusions about the data. Perhaps the most notable of the assumptions is the data should approximately follow a distribution, which is defined and accomplished by using distribution parameter(s). For example the normal distribution is the most commonly selected distribution, which would require the use of tests to check if the data is normality distributed (Hoskin, 2012). The other assumptions for parametric statistics include equal variance, independence, and no outlying observations. These assumptions are confirmed or disproved with the use of parametric tests; however, if these assumptions are not met, they may be able to be reconciled with additional procedures. The assumption of outliers is tackled in Section 2.8.5, as it is a notable concern in this study.

Nonparametric statistics, on the other hand, requires few assumptions about the underlying populations of the data. Namely, the most common assumptions in nonparametric tests are randomness and independence. The parametric assumption of a distribution is not required for nonparametric methods, which are also referred to as distribution-free methods (Hollander and others, 2014). Because a specific distribution is not required, there is a significantly larger number of tests that can be accurately applied to the data. Additional advantages of using nonparametric methods include, but are not limited to: they are generally easier to apply and understand, they are more efficient when the populations are not normal, they can be used with smaller sample sizes, and they are insensitive to outliers (Hollander and others, 2014; Hollander & Sethuraman, 2015).

Based on the requirements and advantages of both statistical procedures, which statistical procedure will be the more effective of the two for this study will be decided upon the completion of descriptive statistics and preliminary analysis of each data set. This will pave the way for the rest of the analysis, indicating which statistical tests should be used (see Section 2.8.6).

2.8.3 Type I Error

In a statistical analysis, Type I Error occurs when the hypothesis test incorrectly rejects the null hypothesis, commonly referred to as “false positives.” Likewise, Type II Error occurs when the test incorrectly fails to reject the null hypothesis, referred to as “false negatives” (Casella & Berger, 2002). The probability of making a Type I Error is represented by the alpha level (α-level) or level of significance, which directly corresponds to the p-value, below which the null hypothesis is rejected (McLeod, 2019). Generally, the default for the α-level is 0.05 ($\alpha = 0.05$), though it is imperative to not automatically revert to this value (Kim, 2015; Maier & Lakens, 2022; Miller & Ulrich, 2019). In order to choose the level of significance, key factors such as sample size and expected losses (false positives and false negatives) must be considered (Kim, 2015).

2.8.4 Nested vs Crossed Designs

A nested design occurs when every level of a factor co-occurs with only one level of another factor, while a crossed design occurs when every level of a factor co-occurs with every level of another factor (Grace-Martin, 2013; Kutner and others, 2004; Stahle & Wold, 1989). Grace-Martin describes this difference by stating the following: “If two factors are crossed, you can calculate an interaction. If they are nested, you cannot because you do not have every combination of one factor along with every combination of the other” (Grace-Martin, 2013).

2.8.5 Outliers

Outliers are extreme observations within the experiment data and can cause serious problems in statistical analysis (Kutner and others, 2004). One possible approach when finding an abnormal observation may be to assume it is the result of an experimental error and discard the data. However, this may not be the correct decision as these possible outliers may impart vital information about the population and study as a whole. The first step in analyzing data for possible outliers is to conduct a thorough examination of residual plots, box plots, stem-and-leaf plots, frequency dot plots, and Grubbs' Tests. Once possible outliers are identified, further research should be conducted to determine the reason for their extreme values. The decision to discard outliers should only be made if there is direct evidence of a recording error, a miscalculation, or equipment malfunctioning (Burke, 1998; Kutner and others, 2004).

As mentioned earlier in this chapter, if the data does not fit a normal distribution with the possible outlier included, there may be cause to consider a nonparametric approach. Nonparametric statistics do not utilize a common assumption that the population is normally distributed and are relatively insensitive to outlying observations (Hollander and others, 2014; Hollander & Sethuraman, 2015). As this study utilizes test scores that we are confident were accurately collected by the professors themselves, there is little reason for discarding possible outliers. Nonetheless, identifying and testing for outliers and analyzing the possible causes of said observations is critical to understanding the data and must be included.

2.8.6 Statistical Tests for Comparing

There is a myriad of statistical tests for comparing two (or more) data sets. In order to properly choose which tests to perform, characteristics of the data sets must be fully discerned first. First, determining the outcomes of interest and whether they are continuous or discrete random variables is required. Then, analyzing the mean, median, range, standard deviations, quartiles, potential outliers, and distribution of a data set can help determine whether or not the assumptions of certain tests are met.

There are various tests, both parametric and nonparametric, available to compare data sets; these include, but are not limited to:

- Independent t-test or paired-samples t-test to compare means of two samples
- Wilcoxon Rank Sum Test to compare medians of two independent samples
- Kruskal-Wallis Test to compare medians (mean ranks) of three or more independent samples
- Conover squared ranks test of equal variance for k samples
- Kolmogorov-Smirnov test to compare the distributions of two samples
- Shapiro-Wilk test to determine if a sample follows the normal distribution
- Chi-squared test of independence for categorical variables

Many of the studies referenced in Section 2.7 utilized at least one of the above tests to compare the grades and demographics of the students in each classroom environment.

2.9 Summary

Understanding instructional technology and distance learning in their entireties and the type of studies that have already been conducted is imperative to further grow this type of research. This enables researchers to determine how student performance is affected by instructional technology and to find possible ways to improve that performance. The next chapter conducts a more specific look at AFIT, which suggests how AFIT may differ from the universities and institutions found in the previous studies. The first data set is then presented and analyzed employing the techniques presented in this chapter.

III. Observational Results and Analysis

3.1 Chapter Overview

This chapter begins by introducing the characteristics that make AFIT unique, along with details and responses from AFIT-EN professor interviews. This gives an idea of the greater population the data is pulled from. The first data set is then further described, and the statistical procedures and processes detailed in Chapter II are employed on that data set. This is followed by model building, and the chapter is concluded with inferences about the model and the data.

3.2 What Makes AFIT Different?

The Air Force Institute of Technology is the Department of the Air Force's leader for advanced, multi-disciplinary academic education, as well as its institution for technical professional continuing education (*About the Air Force Institute of Technology*, n.d.). The Graduate School of Engineering and Management is one of four schools at AFIT and is the "Air Force's advanced academic degree institution which serves the Air Force's STEM workforce needs by providing defense-focused, research-based academic programs leading the award of master's and PhD degrees in engineering, applied science and selected areas of management" (*Fall 2021 Fact Book*, 2022). There are many unique elements and characteristics about AFIT, specifically AFIT-EN, some of which are discussed in the following paragraphs.

Due to the military aspect of AFIT, it is possible that the faculty, staff, and students are atypical of the individuals found at other graduate research institutions. For example, in Fiscal Year 2022, the total faculty at AFIT was composed of 148 civilians and 123 military personnel (*FY2022 AFIT Fact Sheet*, 2022). Generally, military tours are two to four years, which means the turnover rate for faculty is exceptionally high. Not considering the retention rate of civilian faculty, nearly half of all faculty at AFIT will be completely new within four years. In contrast, studies from civilian institutions have shown faculty members stay at a university for a median of 11 years (*Faculty Retention Proves a Major Challenge for Universities*, n.d.). More recent studies during and after the COVID-19 pandemic found that a third of institutions are facing higher student-to-faculty ratios due to movement in their employee base. Additionally, 55% of faculty at higher education institutions have considered changing career or retiring early, including 35% of tenured employees (Umpierrez, 2021). Within AFIT's graduate school, the overall student-to-faculty rate was 4.5 students to every 1 faculty member in Fall 2021 (*Fall 2021 Fact Book*, 2022); while the national average for degree-granting United States institutions was 14-to-1 in 2018 (and may have gotten worse post-COVID), it is unclear how this translates to other graduate schools similar to AFIT (Carlton, 2022). Other universities that are highly focused on research, such as Massachusetts Institute of Technology and California Institute of Technology, had a ratio of 3-to-1 in 2018, which may be a better comparison to AFIT. Therefore, despite the faculty turnover rate at AFIT, students might be getting more focused attention, similar to that of other research institutions.

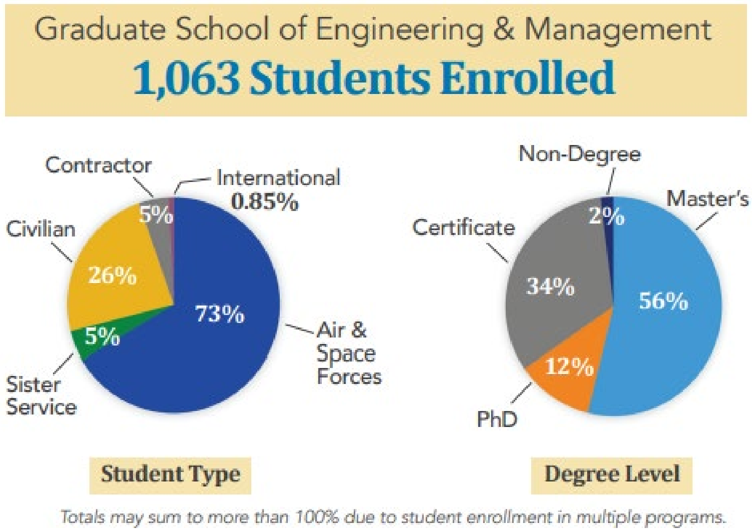


Figure 3: AFIT-EN Student Type and Degree Level Demographics

On a similar note, 78% of students in 2022 were United States Military personnel, with 73% specifically being from the Air and Space Forces. A majority of these individuals (56%) were enrolled in master's programs, while 34% were enrolled in certificate programs, and 12% enrolled in PhD programs, as seen in Figure 3 (*FY2022 AFIT Fact Sheet*, 2022). While some of these military students attend AFIT straight from their undergraduate institutions, many of them have one or two tours their career field (typically 3 to 6 years), outside of an academic environment. Although the statistics on the average student age and family responsibilities aren't available for AFIT, this may provide another contrasting component compared to civilian institutions.

Furthermore, most of these military students that attend courses in-residence don't have another work responsibility. According to the Fall 2021 Fact Book, over 40% of AFIT students were on a quota in Fall 2021 (2022), meaning they were sponsored by various organizations that pay for their tuition. These quota students are bound by a contract that requires them to work for an allotted amount of time (usually after graduation and not necessarily to the same organization that sponsored them) or risk being required to pay the tuition costs back. The sponsorships allow students to focus solely on their school work, without needing to seek further employment to pay personal expenses, since their salary is also paid by the United States Government. Of those students who are not quota, nearly 60% were enrolled in a certificate program, many of whom are DoD civilians or contractors, and may be working another job while attending classes. As a whole of the student population in 2022, DoD civilians and contractors only accounted for 26% and 5%, respectively, as seen in Figure 3. Meanwhile, according to an article by Emily DeRuy and National Journal, 76% of graduate students at civilian institutions work at least 30 hours a week, in addition to their school work, to pay for school and support their families. While working can open career opportunities and offer valuable experience for these students, it is also noted that working long hours can put these students at risk of poor grades or dropping out (DeRuy & National Journal, 2015). The difference in work obligations and hourly time commitments may put AFIT students (particularly quota students) at an advantage over their civilian graduate school counterparts. Ultimately, it is possible that the students at AFIT's graduate school may be a fundamentally different population than their counterparts at civilian institutions, which may determine how students perform with instructional technology.

Lastly, based on interviews conducted with AFIT professors (detailed in the following section) and personal experience, it was acknowledged that the military is often delayed in applying new technology available to the general public. Because of the physical, personnel, and information security requirements, extensive testing is mandatory of any new technology to ensure it meets these requirements before employing it on the military networks. The interviews with AFIT-EN professors indicate this delay puts utilizing technology at AFIT at a severe disadvantage. Outdated software (such as web browsers, SharePoint, digital whiteboards, and student information systems) and outdated hardware (such as computers, monitors, and projectors) leaves professors and students to find their own solutions and tools. The speed, accessibility, and reliability of Wi-Fi on campus is also a major concern when using any kind of technology, whether it be for instruction in the classroom (virtual or in-person classroom), research on personal devices, or communicating with fellow faculty and students. While September 2022 saw a significant upgrade to the AFIT campus Wi-Fi, doubling both the existing wireless footprint and the available bandwidth from 500Mbs to 1Gbs, there are still classrooms that cannot connect to the Wi-Fi and, in times of peak use, virtual classes and meetings suffer in quality and reliability.

The characteristics listed in this section are just a few of the features that make AFIT unique – whether they have positive, negative, or neutral effects on a student's performance remains to be understood. The following section summarizes responses collected from interviews with AFIT professors.

3.3 Professor Interviews

Due to differences in universities and student populations, interviews with two AFIT professors were conducted to gain a better understanding of teaching with instructional technology (IT) at AFIT. The two professors were selected based on having (perceived) differing views on the impact of instructional technology in the classroom and constituted a convenience sample. This section serves as a summary of the detailed responses received from those interviews. The professors who were interviewed were both civilians, which accounted for about 55% of all AFIT faculty in 2022 (*FY2022 AFIT Fact Sheet*, 2022), and have a fair amount of experience teaching with IT, to include teaching 10-30 classes using IT and 4-7 fully distance learning. IT the professors have used includes, but is not limited to: Microsoft Teams (for chat/communications, class meetings, breakout rooms, and feedback from students to instructor), online notebooks (Python, Jupyter, One-Note), SMART boards, the Canvas learning management system (LMS), pre-recorded video lectures, and online textbooks.

Professors were asked about the difficulties and advantages they have experienced while teaching with IT at AFIT. A common difficulty that was mentioned was the internet bandwidth and how the Wi-Fi infrastructure is not sufficient for supporting a hybrid approach in classroom environments. Other difficulties were those mentioned in the previous section that revolved around the technology available at AFIT: outdated and unreliable software and hardware. This was particularly apparent during the COVID-19 pandemic, when professors were forced to find solutions for inadequate technology outside of the AFIT network. Advantages for teaching with IT included: having access to Microsoft Teams (which supports synchronous remote learning), more availability for students (they can refer to recorded lectures and use chat functions which reduced the need for fixed, inflexible, office hours). Moreover, once the lectures have been recorded for one class, they can be stored and easily accessed for future classes. All of these advantages also provide flexibility in professors' and students' personal lives.

When asked what their preferred method of teaching was, there was a mixed response of in-person and synchronous distance learning. When teaching in-person, one professor stated they can better gauge students' reactions – are they sleeping, confused, or grasping the concepts? On the other hand, the other professor preferred synchronous distance learning because the professor could use their own high-performance software and hardware equipment without having to rely on the technology AFIT provides.

The professors were also asked for any recommendations they might have for teachers who are newly integrating IT into their curriculums. Some responses revolved around uncertainty on how to change the mentality of tech-averse faculty to be more tech-eager or even tech-savvy. One suggestion was to update the technology support provided by AFIT to include an assigned videographer to go to classes and set up equipment, record, and edit videos to help encourage professors. For recommendations in designing a course, determining course style or learning environment, along with starting small, were both good places to begin. When trying something new, the professors recommended that asking for feedback can be highly beneficial for improving and fixing things. Suggested questions included:

- What is something you have seen in this class I should keep doing?
- What is something you have seen in this class I should reduce/stop doing?
- What is something from this class you wish other classes would initiate?
- What is something from another class you wish this class had done?

Encouraging students to respond to questions such as these could provide helpful insight into creating a classroom environment that both involves instructional technology and is beneficial to student performance and enjoyment. The topic of feedback is also explored in Chapter V when considering future research.

3.4 Syllabus Summary

This study analyzes the data collected from a four-credit, graduate-level, introductory mathematics course that details vectors, matrices, linear equations, and their applications. The data used is unlike that of other studies previously referenced, but is very indicative of the Air Force Institute of Technology Graduate School of Engineering and Management. Whereas most other studies have one professor for a class spanning one or two semesters that is taught in-person and online (Dutton & Dutton, 2005; Kofoed and others, 2021; Lazari, 2018), this study utilizes data from three professors spanning nine quarters with a mix of in-person and online structures for each professor. As previously stated in Section 2.8, the high turnover of faculty and staff is an important aspect of AFIT; therefore, incorporating multiple professors in the study is necessary to understand how it impacts student performance at the Graduate School.

First, a look at the course similarities is presented. All course sections had the same learning objectives and utilized the same textbook. Also, all sections had homework problems that were suggested to the students to work on and complete, but the homework was not collected or graded. Students were encouraged to work together on these homework problems and solutions were provided. Additionally, all sections had two midterm exams and one final exam. Lastly, all sections, regardless of the learning environment, seemed to follow a linear learning set-up that generally followed the outline from the textbook.

When considering three professors across nine quarters, there were some differences in course structure that will be detailed in the next few paragraphs. These were all identified from the syllabi, but due to the requirement to anonymize the data, to include the quarters and instructors, most were not included in the statistical analysis. Only two sections included quizzes: Quarters 6 and 9. One provided practice quizzes for students and the other had weekly quizzes worth 30% of the total grade (although, due to the anonymization, it is unknown which quarter is which). Similarly, most of the sections had a 30/30/40 split for the total grade utilizing only the three exams, with the exception of two sections: one quarter gave more weight to the final exam and another, as previously stated, allotted 30% to quiz scores in addition to the three exams. There was a mix of online and in-person exams, even among the in-residence and distance learning sections. Furthermore, there were some differences listed in the various syllabi pertaining to the translation from continuous grades to letter grades. These varied between professors and between courses, but each syllabus stated that at the end of the course, when final letter grades were determined, that the thresholds may be lower than identified in the syllabus. Therefore, the relationship between student performance and letter grades should be consistent with tradition at AFIT and academia as a whole. These potential differences in thresholds are considered between professors in Section 3.8.

Of the nine class sections in the data, five were conducted in-residence. As the course was a four-credit course, students were present in lecture for four hours a week in each of the in-residence sections (with the exceptions of any holidays, COVID outbreaks/operational requirements, instructor temporary duty travel (TDY), conferences, etc.).

The remaining four sections were completed via distance learning, with a mix of synchronous, asynchronous, and hybrid learning. Some sections had lectures pre-recorded that students were required to watch at a set pace. In other sections, students had to login to watch the lectures live, as the instructor was presenting. And others had the option of accessing the lectures in real-time or accessing those same recorded lectures at another time convenient for the students.

While the analysis of this type of data has yet to be completed, it is important for this research and for future work to consider as many of these differences as possible in an analysis to understand how these differences do or do not change student performance. When a wider range of sections, teaching styles, and classroom environments are considered and accounted for in an analysis, the impact of instructional technology use may depend less on instructor preference or bias, and more on the actual information comprehension and retention of the students. A summary of these attributes that were not anonymized is displayed in Table 1.

Table 1: Syllabi Summary by Quarter

	Professor	In-Res/DL	Homework	Quizzes
Quarter 1	3	DL	S	N/A
Quarter 2	1	DL	S	N/A
Quarter 3	3	DL	S	N/A
Quarter 4	1	In-Res	S	N/A
Quarter 5	1	In-Res	S	N/A
Quarter 6	2	In-Res	S	Y
Quarter 7	1	In-Res	S	N/A
Quarter 8	3	In-Res	S	N/A
Quarter 9	2	DL	S	Y

Key: S-Suggested: assigned but not collected or graded; solutions provided

3.5 Data Demographics

The data set is comprised of 13 potential predictor variables, and includes continuous data (Exam Grades, undergraduate GPAs, and GRE scores) and nominal and ordinal categorical data. These variables include information about the quarters themselves and demographic information from individual students. The 13 variables are: quarter (randomly and uniquely coded); instructor; in-residence/distance learning (IR/DL); whether quizzes were included; and each student's department, age group, military affiliation, midterm 1 grade, midterm 2 grade, final exam grade, undergraduate GPA, GRE Quantitative Score, and GRE Verbal Score. The response variables are in the form of continuous final grades on a standard scale and letter grades (A-F and W), which was also translated to the 4.0 scale. There are 186 observations in this data set (N=186) and the description of the response and predictor variables are provided in Table 2 and Table 3.

Table 2: Descriptive Statistics of Continuous Response & Variables

Indicator	Variable	Mean	Median	Min/Max	Std. Dev.	25/75 Quartiles
Y_1	Final Grade - Cont	86.4	90	21/105	11.59	82/94
X_1	Midterm 1	89.73	92.5	36.67/102	10.18	85/97.2
X_2	Midterm 2	82.57	88	20/103	17.4	75/94
X_3	Final Exam	87.22	91	23.5/111	13.02	82.25/96.7
X_4	Undergrad GPA	3.31	3.31	2.35/4.0	0.37	3.05/3.58
X_5	GRE Quantitative	157.65	158	139/170	5.52	154/161
X_6	GRE Verbal	155.72	155	140/170	6.02	152/160

Table 3: Descriptive Statistics of Discrete Response & Variables					
Indicator	Variable	Response	Freq	Percent	Mean
Y_2	Final Grade - Letter	A	72	38.71%	95.29
		A-	36	19.35%	89.19
		B+	28	15.05%	83.93
		B	30	16.13%	77.97
		B-	5	2.69%	69.2
		C+	3	1.61%	70
		C	3	1.61%	54.33
		D	3	1.61%	57.33
		F	2	1.08%	36
		W	4	2.15%	---
X_7	IR/DL	0 (IR)	93	50.00%	84.5
		1 (DL)	93	50.00%	88.33
X_8	Professor	A	72	38.71%	79.66
		B	42	22.58%	90.44
		C	72	28.71%	90.7
X_9	Quarter	1	25	13.44%	88.17
		2	12	6.45%	74.91
		3	23	12.37%	93.7
		4	21	11.29%	82.05
		5	23	12.37%	82.87
		6	9	4.84%	94.78
		7	16	8.60%	75.31
		8	24	12.90%	90.38
		9	33	17.74%	89.22
X_{10}	Department	ENG	19	10.22%	83.72
		ENO	16	8.60%	81.87
		ENP	5	2.69%	95.2
		ENS	60	32.23%	85.76
		ENV	2	1.08%	73.5
		ENY	84	45.16%	88
X_{11}	Age	20-29	142	79.33%	87.59
		30-39	27	15.08%	83.81
		40+	10	5.59%	79.43
X_{12}	Military Affiliation	0 (Mil)	149	80.11%	86.14
		1 (Civ)	37	19.89%	87.53
X_{13}	Quizzes Included	0 (No)	144	77.42%	85.22
		1 (Yes)	42	22.58%	90.44

Since this study is focused on how learning environments affect students, the demographics for the predictor variables is included contingent on the IR/DL status of the students. These statistics are found in Table 4 and Table 5, with the exception of

X_9: Quarter and X_{13}: Quizzes Included since that information was presented in Table 1. Table 4 presents the continuous variables and the mean of each in relation to the IR/DL status. Table 5 presents the categorical variables with the frequency of each group in relation to the IR/DL status, along with the percentage of that frequency within the variable's specific group.

Table 4: Continuous Variable Break-out Contingent on X_7: IR/DL

Indicator	Variable	Course Status	Mean
Y_1	Final Grade - Cont	IR	84.5
		DL	88.33
X_1	Midterm 1	IR	88.15
		DL	91.35
X_2	Midterm 2	IR	75.8
		DL	89.57
X_3	Final Exam	IR	85.87
		DL	88.63
X_4	Undergrad GPA	IR	3.28
		DL	3.33
X_5	GRE Quantitative	IR	157.84
		DL	157.35
X_6	GRE Verbal	IR	156.24
		DL	154.9

Table 5: Discrete Variable Break-out Contingent on X_7: IR/DL					
Indicator	Variable	Response	Course Status	Freq	Percent
X_8	Professor	A	IR	60	83.33%
			DL	12	16.67%
		B	IR	9	21.43%
			DL	33	78.57%
		C	IR	24	33.33%
			DL	48	66.67%
X_{10}	Department	ENG	IR	15	78.95%
			DL	4	21.05%
		ENO	IR	9	56.25%
			DL	7	43.75%
		ENP	IR	2	40.00%
			DL	3	60.00%
		ENS	IR	13	32.23%
			DL	47	95.92%
		ENV	IR	2	100.00%
			DL	0	0.00%
		ENY	IR	52	61.90%
			DL	32	38.10%
X_{11}	Age	20-29	IR	67	47.18%
			DL	75	52.82%
		30-39	IR	16	59.26%
			DL	11	40.74%
		40+	IR	3	30.00%
			DL	7	70.00%
X_{12}	Military Affiliation	0 (Mil)	IR	73	48.99%
			DL	76	51.01%
		1 (Civ)	IR	20	54.05%
			DL	17	45.95%

Using the demographics for this data, a comparison can be made with overall AFIT demographic statistics. These comparisons are done utilizing the Fall 2021 Fact Book and Application Requirements from the AFIT-EN Office of Admissions (*AFIT / Office of Admissions / Air and Space Force*, n.d.; *Fall 2021 Fact Book*, 2022). This will show how representative of a sample this class is of the general AFIT population. Based on the demographic information received in this data set, military affiliation, department, and program/degree type are compared in Table 6.

Table 6: Demographic Information Comparison A		
	Overall AFIT	Study
Military Affiliation		
Military	72.07%	80.11%
Civilian	27.93%	19.89%
Department		
ENC	1.58%	0%
ENG	15.10%	10.22%
ENO	2.16%	8.60%
ENP	16.85%	2.69%
ENS	28.13%	32.23%
ENV	23.65%	1.08%
ENY	12.53%	45.16%
Program		
Certificate	33.95%	25.81%
Masters	53.32%	62.90%
PhD	10.78%	2.15%
Non-Degree	1.95%	9.14%

For the study and analysis of the data set in this chapter, a Type I Error or α-level of 0.05 was chosen based on a large original sample size (186 observations) and the social sciences, human factors and educational aspects. In choosing the α-level based on these factors specific to this study, statistical inferences should be improved and the efficiency and informativeness of the research may increase (Maier & Lakens, 2022). Pearson Chi-Squared (χ^2) tests were conducted to determine significant differences between the frequencies of this study and overall AFIT-EN within each group. The χ^2 test for Military Affiliation resulted in a p-value of 0.0146, indicating there is a significant statistical difference between this study and AFIT-EN. A visual comparison shows there some differences in department, which may be attributed to specific program requirements and course prerequisites. In fact, the χ^2 test for Department confirmed these visual differences with a p-value result of <0.0001. Lastly, there are also some differences in program/degree type that can be seen visually but can be attributed to the fact that this is an introductory course, and therefore not many PhD students would need or want to take the course. The

χ^2 test for Program again resulted in a p-value of <0.0001, indicating there is a significant statistical difference in the probability of a student's program or degree type between this study and the overall AFIT-EN population. Based on the Pearson Chi-Squared tests conducted, the final conclusion is that based on the demographic information in Table 6 for Military Affiliation, Department, and Program, there are significant differences in the population sample of this study and the overall AFIT-EN population. However, because the data for this study came from one specific class within one department, these differences are to be expected as this is an introductory-level class, several departments have this class listed as a program requirement and this class is listed as a prerequisite for other classes. These differences may limit the applicability of the results presented here to AFIT-EN courses similar to this introductory course (compared to courses at large), and indicates an area for future research.

Further comparisons between the overall AFIT-EN population and the demographics for this study are provided in Table 7.

Table 7: Demographic Information Comparison B

	Overall AFIT	Study
Average Class Size	14.8	20.7
Average Class Size - ENC	18.8	20.7
Undergrad GPA*	3.00	2.35/3.31
GRE Verbal*	153 (500)	140/155.72
GRE Quantitative*	148 (600)	139/157.65

Key: * - stated minimums for AFIT-EN enrollment (requirements vary); minimums/means provided for this study

While this study does not seem to be a good representative of the average class size of all AFIT-EN classes, it is very close to the average class size for ENC, the Department of Mathematics and Statistics where the course is taught. Standard deviations would assist in further ascertaining the magnitude of this difference, but this information was not available; therefore, a qualitative assessment of the results is provided. For example, the difference of 1.9 in class size can once again be attributed to more students enrolling in introductory classes rather than higher doctorate-level classes. The undergrad GPA, GRE Verbal, and GRE Quantitative requirements were provided by the AFIT-EN Office of Admissions: "the requirements vary, but generally speaking, a grade point average (GPA) of 3.0 on a 4.0 point scale is needed…, as well as either a Graduate Record Examination (GRE) score of 153 (500) verbal/ 148 (600) quantitative…" (*AFIT / Office of Admissions / Air and Space Force*, n.d.). From this wording, it can be deduced that while those are the stated requirements, they are neither the bare minimums for acceptance into AFIT nor the averages for those accepted. The 25% quartiles provided in Table 2 for this study may be more accurate for comparisons. Overall, the conclusion is made that this data set is a representative sample of the overall AFIT population and further analysis and conclusions based on this data can be made for AFIT-EN as a whole.

Lastly, a more thorough analysis on the continuous final grade (Y_1) is conducted. Figure 4 shows the distribution of Y_1 as a histogram and graphed on a normal quantile plot. The response does not seem to follow the normal distribution and a Shapiro-Wilk test rejects the null hypothesis that the data is from the normal distribution (p-value < 0.0001) Furthermore, the data does not follow any other common distributions (lognormal, Weibull, gamma, beta, etc.). Additionally, the graph in Figure 4 indicates there may be issues with at least one outlier, and some tests indicate there may be two – the two lowest final grades. Although modeling may explain and eliminate these issues, a combination of parametric and nonparametric approaches will be utilized for statistical analysis in this data set, depending on the subset of data utilized for each test.

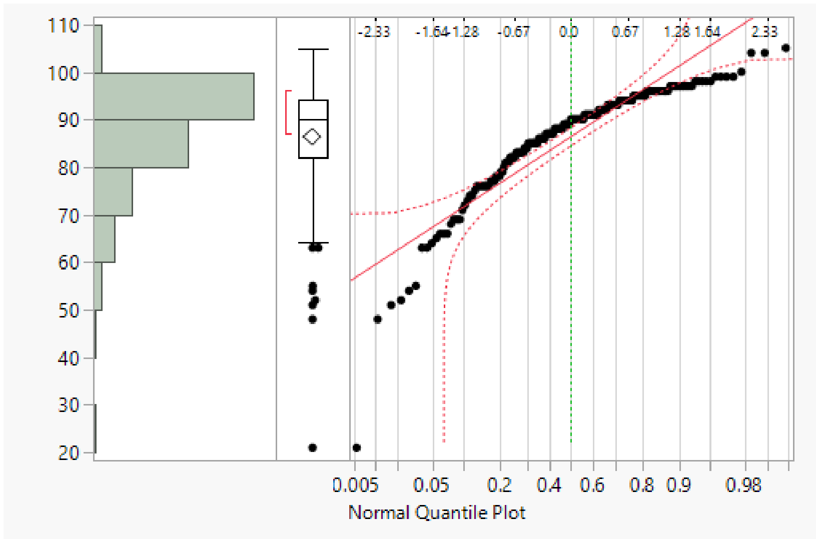


Figure 4: Normal Quantile Plot of Y_1: Final Grade – Continuous I

3.6 Withdrawals and Missing Data

Before analyzing the data collected, it is imperative to understand the missing information within the data. First, there were four students who were annotated as student withdrawals, meaning at some point the student decided to drop the class. These four students were not included in the analysis for two reasons. First, there was no way to tell when the students dropped, and therefore no way to tell if all students who were enrolled and decided to withdraw were included in the data set (that is, including students who may have dropped the class before the add/drop period ended). Second, there was no way to tell why these students withdrew from the class. It could be because they were struggling with the information, found they didn't have the time to dedicate to the class, had personal matters to contend with, etc. Regardless of the reason, a systematic assumption can be made that the students didn't drop based on the differences between each class (e.g., professor, fall or spring quarter, etc.).

Like many studies, missing data is also a potential cause for concern. While there were no missing grades or missing information from the courses themselves, some of the student demographic information was absent. This could have been for a few reasons: the student may not have been able to be identified due to a common name or wasn't in the electronic system. Also, as mentioned in Chapter I, AFIT was included in the schools that temporarily suspended the requirement for GRE scores during the COVID-19 pandemic. This is likely the cause for most of the missing GRE scores in this data set. Once again, a systematic assumption can be made that this missing data is random and does not have to do with how well a student performed academically in the class. Missing data was present in 47.85% of the observations; with some overlapping observations, 18.82% of GPA scores and 33.87% of both GRE Verbal and GRE Quantitative scores were missing. Imputing the missing data field was explored but with the percentage missing, any analysis and conclusions could only be used for hypothesis testing (Dong & Peng, 2013; Madley-Dowd and others, 2019). For this reason, any observations with missing data were omitted when handling those variables. Since the data is believed to be missing at random, this choice will not change the population about which inference can be made. A summary of the missing observations and final sample sizes is provided in Table 8.

Table 8: Observations & Sample Sizes

Variable	Initial n	Missing	Final n
Age Group	186	7	179
Undergrad GPA	186	35	151
GRE Quantitative Score	186	63	123
GRE Verbal Score	186	63	123
Model Building	186	90	96

3.7 Multivariate Correlations

Next, we consider the correlations amongst the response and predictor variables and within the predictor variables to understand which interactions might be most significant. The only notable correlations were found between the Final Grade – Continuous (Y_1) and Midterm 2 (X_2) and Final Exam (X_3) with correlations of 0.8395 and 0.9005, respectively. This is to be expected since the exam grades are percentages of the final grade, but it is interesting to note that Midterm 1 (X_1) does not have the same degree of correlation (0.5885). One possible reason for this difference in correlation could be that students didn't know the expectation for the class and exams and therefore didn't know how to prepare for the exams until after the first midterm was conducted. Students then either decided to drop the course altogether or to put in extra work as a consequence of receiving a low midterm grade to ensure their subsequent test grades and final grade were satisfactory. The next highest correlation was between Quarter (X_6) and Quizzes Included (X_{14}) with a correlation of 0.6102, which indicates there is very low to no correlation between independent variables. Since there were only two quarters that had quizzes, and no students in other quarters had quizzes, this degree of correlation between the two variables is also to be expected. There were no other correlations for variables outside of the exam grades with absolute values higher than 0.4, and therefore are not deemed significant to highlight or include in further analysis. The scatterplot matrix in Appendix B displays most of the main factors, including the response, plotted against each other.

3.8 Factor Level Analysis

Subsequently, a thorough analysis is conducted on how the groupings within each categorical variable relates to students' final grades. Most of the testing in this section was conducted using a nonparametric approach since not all the samples and groups were normally distributed. A summary of the tests utilized and the results discussed in this section is displayed in Table 9.

Table 9: Summary of Statistical Tests & Results I

Response x Variable	Type of Test	Test	P-value	Result
Final Grade x X_7: IR/DL	Nonparametric	Levene	0.2025	Equal variances
	Nonparametric	Kolmogorov Smirnov	0.0059	Not from same distribution
	Nonparametric	Wilcoxon Rank Sum	0.0040	Unequal means (DL higher than IR)
Final Grade x X_8: Instructor	Nonparametric	Levene	0.0193	Unequal variances
	Nonparametric	Kruskal-Wallis	< 0.0001	Unequal medians (A lower than B & C)
Final Grade x X_{10}: Department	Nonparametric	Levene	< 0.0001	Unequal variances
	Nonparametric	Kruskal-Wallis	0.0907	Equal medians
Final Grade x X_{11}: Age	Nonparametric	Conover Squared Rank	0.0608	Equal variances
	Nonparametric	Kruskal-Wallis	0.1853	Equal medians
Final Grade x X_{12}: Military Affiliation	Nonparametric	Levene	0.7956	Equal variances
	Nonparametric	Kolmogorov Smirnov	0.9526	Same distribution
	Nonparametric	Wilcoxon Rank Sum	0.4383	Equal means

Figure 5 shows a box plot of the final grades in relation to in-residence and distance learning sections. The mean for the final grade in in-residence sections is 84.5, while the mean for the final grade in distance learning sections is 88.33 (as shown in Table 3 and the X's within the box in Figure 5). While it is clear the distance learning average is higher than that of the in-residence, further testing must be completed to determine if the difference is statistically significant. Since the Levene variance tests concluded the variances are equal (p-value = 0.2025), a Kolmogorov-Smirnov test was used, which concluded the two populations were not from the same distribution (p-value = 0.0059). Since the Wilcoxon Rank Sum Test is nonparametric and distribution free, it is utilized to compare these two means, resulting in a p-value of 0.004. As this is less than our chosen significance level of $\alpha = 0.05$, we reject the null hypothesis and conclude the mean of the distance learning sections is statistically higher than the in-residence sections.

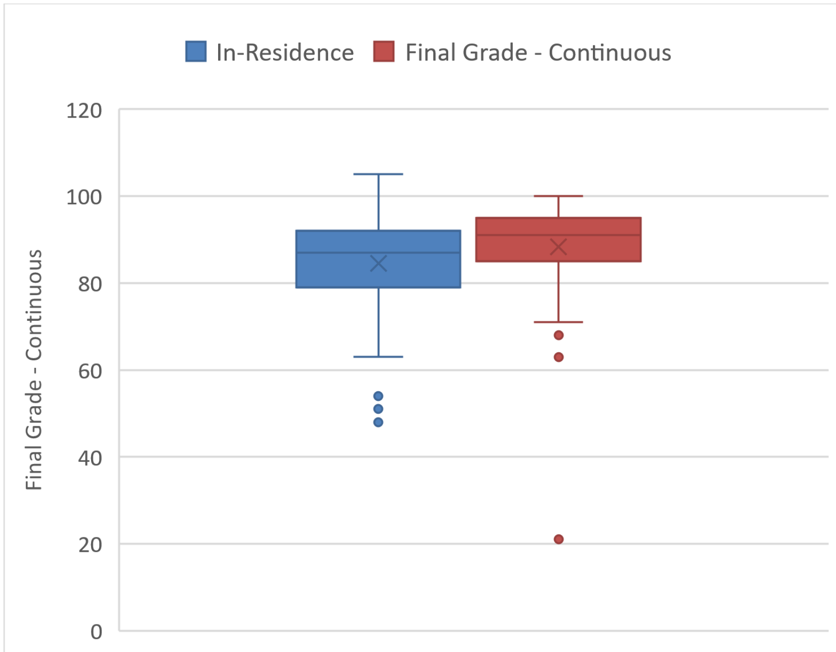


Figure 5: Box Plot of Final Grade x IR/DL

In Table 3, the final grade means of the nominal variables X_8: Professor, X_{10}: Department, X_{11}: Age, and X_{12}: Military Affiliation were also broken out based on the final continuous grade response. The same process used for comparing X_7: IR/DL to the response will be utilized to consider how these four variables each affect the response, Y_1: Final Grade – Continuous.

Regarding X_8, the Levene test for variance concluded all three instructors did not have the same variance in their final grades. However, a visual inspection displayed the distributions were similar enough to conduct the Kruskal-Wallis test. This test indicated at least one professor had final grades that were statistically significantly different than the other two (p-value < 0.0001). Further review utilizing the pairwise Wilcoxon rank sum test with alpha adjustments showed instructor A had significantly lower mean scores than instructors B and C. Recall that Section 3.3 introduced potential differences in letter grades translations and thresholds. Figure 6 shows this difference in instructor grading in terms of Y_2: Final Grade – Letter. Despite the idea that instructors may have different conversions from continuous grades to letter grades (thereby making the letter grades consistent across AFIT-EN as a whole), there still appears to be a significant disparity between the grades obtained from instructor A and instructors B and C.

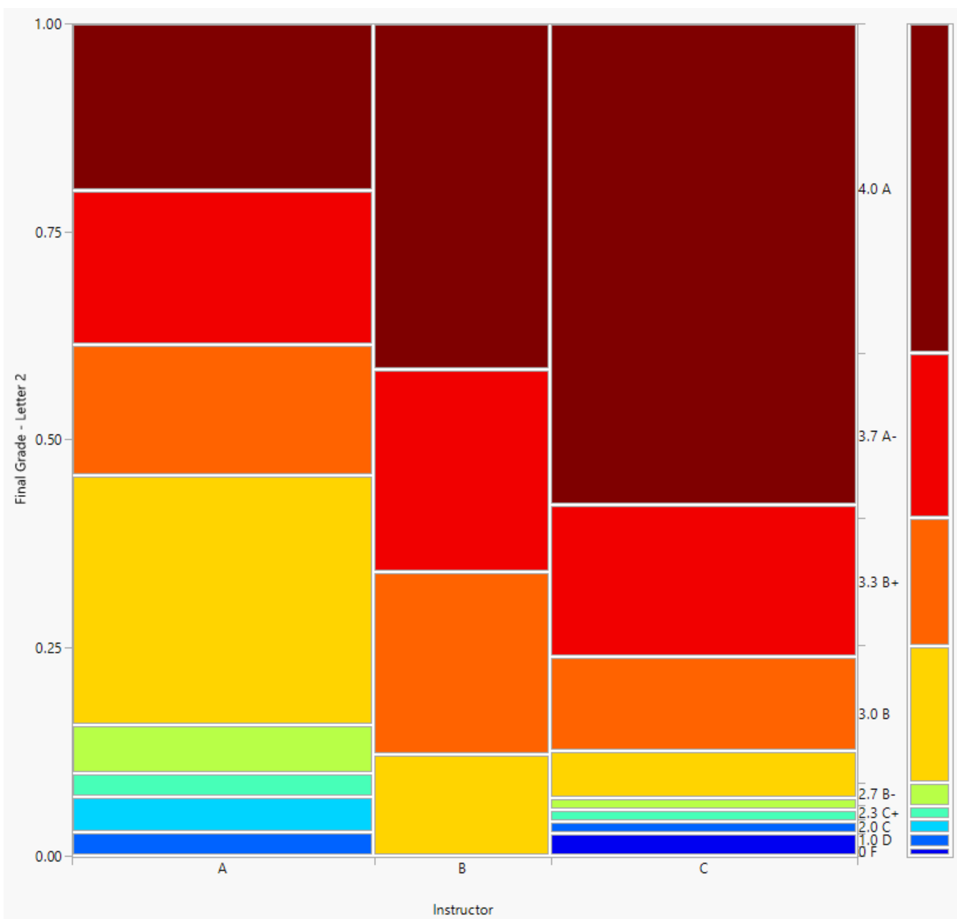


Figure 6: Frequency Plot of Y_2: Final Grade – Letter by Instructor

Next, consider X_{10}: Department and its impact on the Y_1: Final Grade – Continuous. Variance testing indicated there was significant differences between the variances of the groups (Levene test with p-value < 0.0001). Because of this, robust testing to account for unequal variances was utilized, which indicated there was no significant differences between the medians of the groups (Kruskal-Wallis with p-value = 0.0907). However, these conclusions may not be very robust as the sample sizes for two of the departments were very small, with ENV only having 2 observations and ENP only having 5 observations. Indeed, when utilizing the pairwise Wilcoxon rank sum test with alpha adjustments, the results indicated there was significant differences between ENP and both ENS and ENY.

Furthermore, the Conover Squared Rank test demonstrated there was no significant difference in variances between X_{11}: Age and Y_1: Final Grade – Continuous (p-value = 0.0608). The Kruskal-Wallis test for medians also indicated there was no significant difference (p-value = 0.1853).

Lastly, the relationship between X_{12}: Military Affiliation and the response Y_1: Final Grade – Continuous is evaluated. Here, there is once again only two groups in the variable and the Levene test indicated no significant difference with a p-value of 0.7956. Likewise, the Wilcoxon Rank Sum Test showed no significant difference in the means of the two groups (p-value = 0.4394). The Kolmogorov Smirnov test solidified the results of the previous two tests and demonstrated the two groups within military affiliation are from the same distribution with a p-value of 0.9526.

All of the factor level analysis conducted signifies that model building may be successful, given that multiple variables were shown to significantly impact the response, Final Grade – Continuous.

3.9 Model Building and Selection

For model building, a parametric approach was considered and utilized. Additionally, as previously shown in Table 8 due to the missing data, there are 96 total observations for this process. A full model was developed using variables X_4 through X_{13}. Since the final grade is a direct calculation of the examination grades (and quiz grades for one class), these variables were not used for predictive modeling.

Two factors under consideration that require special consideration are quarter and instructor for this data set. Here, quarter is nested in instructor, since each quarter only has one instructor but each instructor taught during multiple quarters. A variable titled "Quarter (random)" was created that was explicitly nested in instructor by giving each quarter a unique random code of one through nine (1-9). By doing this, instead of leaving the quarters implicitly nested, we were able to avoid some linear dependencies and other issues with model building.

Upon further testing, it was found that some variables and interaction terms had linear dependencies with each other; therefore, some of the variables were removed, based on which ones were least significant. The factor level analysis in the previous section also helped to determine which variables were least significant, leading to the removal of age and military affiliation.

A mixed, step-wise approach was used along with sequential tests using the chosen $\alpha = 0.05$ to determine which variables were statistically significant to produce the following reduced model:

$$Y_i = \beta_0 + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \beta_8 X_{8i} + \beta_{10} X_{10i} + \beta_{46} X_{4i} X_{6i} + \varepsilon_i$$

Where: Y_i = Final Grade – Continuous, X_4= Undergraduate GPA, X_5= GRE Quantitative Score, X_6= GRE Verbal Score, X_7= IR/DL Status, X_8= Professor, and X_{10}= Department.

The effects coding for the nominal variables were as follows:

X_7= {1 for IR (0), -1 for DL (1)}
X_8= {1 for A, -1 for B & C}
X_{10}= {1 for ENS & ENG, -1 for ENY and ENP, 0 for all else}

There were many interactions considered in the full model, but the interaction between Undergraduate GPA and GRE Verbal Score was the only one that presented significant evidence of having an influence in the reduced model.

Once the reduced model was created, the diagnostics to evaluate model assumptions were inspected. Normality and constant variance were absent; in attempts to address these violations, a Box-Cox transformation was conducted, resulting in $\lambda = 4.65$. The software utilized (JMP) employs the following formula to conduct a Box-Cox transformation (where $\tilde{y}$ indicates the geometric mean):

$$Y_\lambda = \begin{cases} \dfrac{y^\lambda - 1}{\lambda \tilde{y}^{\lambda-1}} \; if \; \lambda \neq 0 \\ \tilde{y} \ln y \; \; if \; \lambda = 0 \end{cases}$$

Upon implementing this transformation, there was no evidence of non-constant variance nor non-normality. However, the Durbin-Watson test identified that there was an issue with independence of the residuals (i.e., that autocorrelation existed) with a p-value of 0.0001. Because of this issue, further remedial measures were required.

Since this data set is comprised of individuals' sensitive information and needed to be deidentified, the addition of a variable with time-ordered effects on the response (e.g., a quarter variable that was ordered by time, instead of randomly ordered) was not a possible remedy. Therefore, the Cochrane-Orcutt procedure was utilized to fix the autocorrelation concern. First, an estimation of ρ is accomplished by fitting a straight line through the origin of a model fitting the residuals ε_t to ε_{t-1}; where ε_t and ε_{t-1} are the residuals from the model previously introduced after the Box-Cox transformation was applied in which $t-1$ represents the ordered observation prior to the t^{th} observation. The slope of this straight line with the residuals and through the origin is then annotated as $\hat{\rho}$:

$$\varepsilon_t = (\varepsilon_{t-1})\,\hat{\rho}$$

In this study, it was found that $\hat{\rho} = 0.3517042$. The response and continuous variables are then transformed by applying the following equations: $Y'_t = Y_t - \hat{\rho}\,Y_{t-1}$ (where Y_t and Y_{t-1} are the transformed responses that resulted from the Box-Cox transformation) and $X'_{it} = X_{it} - \hat{\rho}\,X_{i(t-1)}$. The new response and variables are then refit in the model, producing the new reduced model:

$$Y'_{it} = \beta_0 + \beta_4 X'_{4it} + \beta_5 X'_{5it} + \beta_6 X'_{6it} + \beta_7 X_{7i} + \beta_8 X_{8i} + \beta_{10} X_{10i} + \beta_{46} X'_{4it} X'_{6it} + \varepsilon_i$$

Upon completion of this procedure, the Durbin Watson test was recalculated and the new p-value for the test was 0.2244, confirming there is no serial correlation in the new reduced model.

Since a new model was selected, the rest of the model assumptions must be re-confirmed, or remedied if they are found lacking. Normality of the residuals was checked with a normal quantile plot (Figure 7) and a Shapiro-Wilk Test confirms this assumption with a p-value of 0.8782. A plot of the residuals against the predicted values for Y, found in Figure 8, confirms this model has constant variance. This plot also confirms non-linearity is not an issue. Lastly, along with the aforementioned figures, the studentized residuals and the Cook's Distance estimate confirm there are no observations which may be deemed as outliers. All of these conclusions validate the appropriateness of this model.

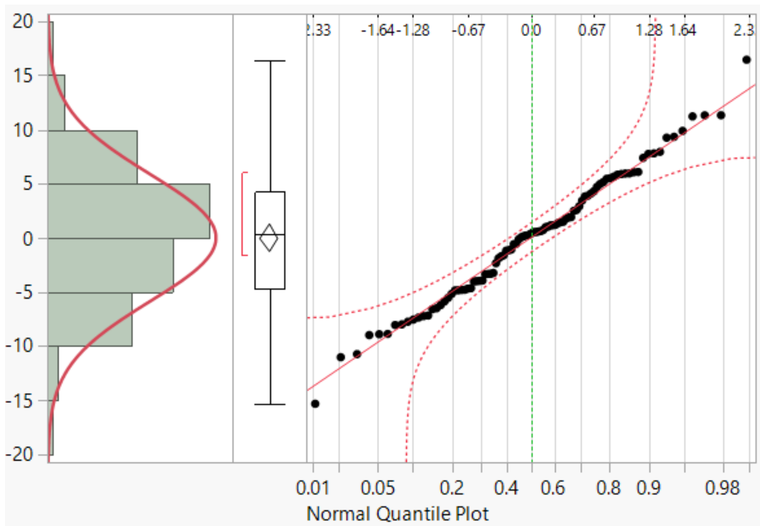


Figure 7: Normal Quantile Plot of Residuals I

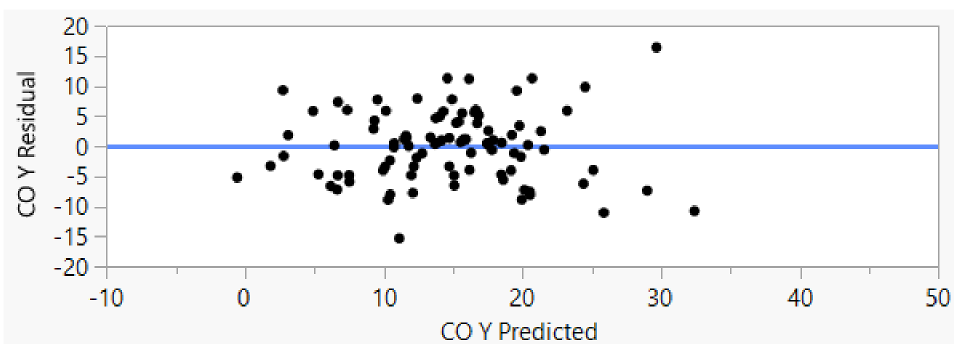


Figure 8: Plot of Y'_t Predicted x Y'_t Residual I

Next a look at the parameter estimates, summary of fit, and analysis of variance of this model is completed to understand the predictiveness of the model. After both transformations, the parameter estimates and details are found in Table 10.

Table 10: Parameter Estimates I				
Term	Estimate	Standard Error	t Ratio	Prob > |t|
Intercept	-29.4413	12.48471	-2.36	0.0206
Dept{ENS&ENG-ENY&ENP}	-1.402975	0.693965	-2.02	0.0463
Instructor{A-C&B}	-4.223522	0.790085	-5.35	< .0001
CO GPA	6.5463727	1.867008	3.51	0.0007
CO GRE Q	0.3972802	0.115946	3.43	0.0009
CO GRE V	-0.129087	0.104703	-1.23	0.2209
(CO GPA-2.13402)*(CO GRE V-100.858)	1.2043175	0.262405	4.59	< .0001
IR/DL[0]	1.3892468	0.714452	1.94	0.0551

All p-values are under the chosen alpha-level of 0.05 with the exception of two. The Cochrane-Orcutt (CO) GRE Verbal was kept in the model since the interaction between the GPA and GRE Verbal was found to be highly significant. To understand this interaction term in the model, an interaction plot was created and is displayed in Figure 9. While it is difficult to interpret the actual numbers in the plot since they are from a twice-transformed space, the relationship between the variables and the response can be interpreted. For example, consider what happens in the lower left quadrant where CO GPA is the X-axis and CO Y is the Y-axis. The blue line indicates a high CO GRE Verbal value of 114.54 and has a positive relationship between GPA and the response. Conversely, the red line indicates a low CO GRE Verbal value of 83.211 and has a negative relationship between GPA and the response. Therefore, a high GPA does not necessarily correlate with a high final grade (CO Y); a high GRE Verbal score is also desired. The upper right quadrant demonstrations the same relationship with CO GRE Verbal on the X-axis and a high and low value of CO GPA indicated by the blue and red lines, respectively. By including the interaction term between undergraduate GPA and GRE Verbal in the model, we can capture relationships that change based on the value of another variable (Frost, 2017b).

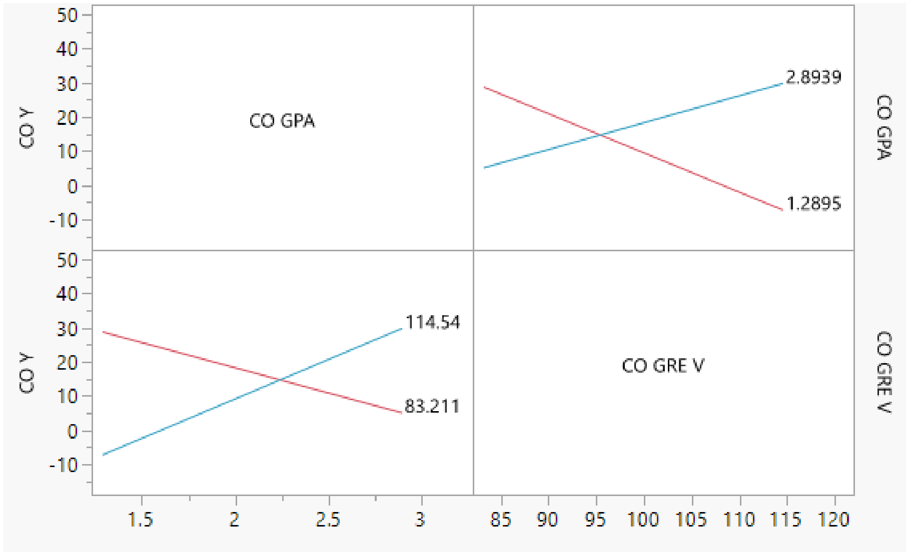


*Figure 9: Interaction Plot of CO GPA*CO GRE Verbal*

Additionally, the In-Residence/Distance Learning variable is the other variable that had a p-value higher than 0.05; this variable was kept in the model since it is the focus of this study. While it was found to be significant after the Box Cox transformation, the changes wrought after the Cochrane-Orcutt procedure deemed the variable not statistically significant to the response. Furthermore, the effect sizes for the parameters are provided in Table 11.

Table 11: Model Effect Sizes

Term	DF	Sum of Squares	F Ratio	Prob > F	η^2	$\eta^2_{partial}$	ω^2	Cohen's d
Dept{ENS&ENG-ENY&ENP}	1	152.8601	4.0872	0.0463	0.0222	0.04487	0.01668	0.064
Instructor{A-C&B}	1	1068.737	28.576	<.0001	0.15522	0.24725	0.14898	0
CO GPA	1	459.8101	12.2945	0.0007	0.06678	0.12382	0.06102	0.002
CO GRE Q	1	439.087	11.7404	0.0009	0.06377	0.1189	0.05802	0.002
CO GRE V	1	56.8478	1.52	0.2209	0.00826	0.01717	0.00281	0.221
CO GPA*CO GRE V	1	787.7842	21.0639	<.0001	0.11441	0.19492	0.10839	0
IR/DL	1	141.4108	3.7811	0.0551	0.02054	0.04165	0.01502	0.064

The summary of fit and analysis of variance details are found in Figures 10 and 11. The coefficient of determination (R-squared), which is a statistical measure that indicates how much variation in the response (Y) is explained by the dependent variables in the model, is provided in Figure 10. An adjusted R-squared value of 0.4894 makes this model on par with industry expectations for a study pertaining to human factors and human behavior, which tend to have R-squared values less than 0.5 (Ballard, 2019; Frost, 2017a). The overall F Ratio and Prob > F (p-value) from Figure 11 are further evidence of the significance of this model as a whole. Finally, due to random missing data and the loss of the first observation from the Cochrane-Orcutt procedure, this model utilized 95 of the original 186 observations. It is possible that had all 186 observations been included, model building may have had a different outcome.

Summary of Fit

RSquare	0.527443
RSquare Adj	0.489421
Root Mean Square Error	6.115538
Mean of Response	14.6149
Observations (or Sum Wgts)	95

Figure 10: Summary of Fit I

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	7	3631.7016	518.815	13.8721
Error	87	3253.7833	37.400	**Prob > F**
C. Total	94	6885.4849		<.0001*

Figure 11: Analysis of Variance I

3.10 Inferences

The principal purpose of model building is to use the capacity and capability of the model to make predictions and inferences on the data. For this study, various factor levels were input into the model to determine the prediction intervals of potential students' final grades. These inputs and prediction intervals are displayed in Table 12.

Table 12: Sample Prediction Intervals									
Dept	Undergrad GPA	GRE Quantitative	GRE Verbal	Instructor	IR/DL	Y'_tp	C-O 95% PI	B-C 95% PI	Original 95% PI
ENY	3.31	158	155	C	IR	20.7002	[8.2376, 33.1629]	[8.8834, 33.8087]	[72.9868, 97.2908]
ENY	3.31	158	155	C	DL	17.9218	[5.5135, 30.32998]	[6.1593, 30.9758]	[67.4591, 95.477]
ENO	3	165	140	A	IR	16.2673	[3.205, 29.3297]	[3.8507, 29.9755]	[58.6175, 94.8053]
ENO	3	165	140	A	DL	13.4889	[0.2184, 26.7593]	[0.8642, 27.4051]	[44.2198, 92.995]
ENS	3.6	150	160	B	IR	18.6108	[5.8495,31.3721]	[6.4953, 32.0179]	[68.234, 96.1588]
ENS	3.6	150	160	B	DL	15.8323	[3.1158, 28.5489]	[9.6111, 29.1947]	[74.2331, 94.2687]
ENO	2.7	143	140	A	IR	9.2111	[-3.8305, 22.2526]	[-3.3413, 22.7418]	[60.9088, 89.3386]
ENO	2.7	143	140	A	DL	6.4326	[-6.8108, 19.6759]	[-6.3216, 20.1651]	[67.8374, 87.0579]
ENP	3.95	170	162	B	IR	38.4138	[24.8968, 51.9308]	[25.386, 52.42]	[91.477, 106.9136]
ENP	3.95	170	162	B	DL	35.6353	[22.1847, 49.086]	[22.6739, 49.5752]	[89.2812, 105.6384]

Each prediction interval was calculated in the transformed space and calculations were completed to return the interval to the original space of the response – a final continuous grade on a standard scale. A variety of data attributes were chosen to display the variations in the model and to indicate how the response changes with the changing variables. While some of the prediction intervals presented do span a wide range of grades that is considered by AFIT standards as passing and failing, this is to be expected due to the model's R-squared, the parameters' standard errors, and the human factors field of study. Additional variables that were not provided in this study may increase the predictive capabilities of a model, which is further discussed in Chapter V.

Despite the factor level analysis from Section 3.8 concluding distance learning students performed better than in-residence students, the model and prediction intervals when using the reduced data set and considering other variables, such as instructors, illustrate in-residence students slightly out-performed distance learning. However, the parameter estimate and p-value found in Table 10 indicated this difference in classroom environments was not statistically significant.

3.11 Conclusion

In considering the final model and the effects, the conclusion may be made that whether a student is enrolled in the in-residence or distance learning section of this mathematics course does not have a significant impact on a student's performance, as exhibited by final grade. The following chapter describes and evaluates the second half of the data set from an experimental study conducted on the same introductory mathematics course.

IV. Experimental Results and Analysis

4.1 Chapter Overview

This chapter presents an experimental study that was conducted on the same four-credit, graduate-level, introductory mathematics course as the previous chapter's data set. The background and setup of the study, along with the key elements of this course, are detailed, followed by an in-depth analysis of the data collected. Finally, conclusions are made in regard to this experimental data and compared to the data from Chapter III.

4.2 Background

The second data set of this study, which is analyzed in this chapter, is the result of an experimental study that was designed in collaboration with the professor of the class, who is a different professor than the three professors from the first data set. This distinction ensures the first data set and second data set are reasonably independent of each other. There are two sections of the class that were included, one in-residence and one distance learning.

As previously mentioned, experimental studies that deal with education and student performance are not common. Certain factors of the class utilized in this study were able to be manipulated, which are detailed in the following section, potentially making the associations and conclusions from this data set to be stronger than the first data set. However, there was one significant aspect of the class that was not able to be controlled: students were not able to be randomized in each section. Students chose whether to enroll in the in-person or distance-learning course. This was the case for all but one of the studies in Chapter II: the experimental study conducted by Kofoed and others randomized the students in each section of the course (2021).

In addition to obtaining IRB Exemption approval, on the first day of class, the professor announced to the students that the class would be a part of an experimental study in which their demographic data and grades would be collected and utilized. This information was also included in the class syllabus. If they wished not to participate, they could disclose their decision to the professor and their information would not be included in any portion of the study; however, all students decided to participate. As with the data from the previous chapter, all the data was deidentified prior to the author receiving it. To maximize student participation, no names were provided by the professor, and therefore couldn't be given to Institutional Research to collect additional demographic data (such as the undergraduate GPA, GRE scores, and age that were included in the first data set). Additionally, once the study and key elements of the class were designed, there was no correspondence between the professor and author to ensure objective results.

4.3 Key Elements

There are four key elements in these two course sections that were included to understand the impact of using technology on student performance. These four elements were suggested to better control for other variables (besides in-residence and distance learning) that were believed to increase student performance, found both in the studies from the literature review and in the preliminary analysis of the first data set course sections in this study. As such, these four elements have been incorporated into both course sections to explore their impact on student performance through a more direct comparison than in Chapter III.

1. *Practice quizzes.* Students were given access to online practice quizzes that were designed to aid in students' understanding and prepare them for the examinations. These practice quizzes were not formally graded and solutions to each quiz were provided. Likewise, practice homework problems were provided, and also not collected or graded.
2. *Exam format.* All students were given all exams in the format they had for class. This means in-residence students took all midterms and the final in-person, all at the same time. The distance learning students took the same midterms and final as their in-residence counterparts, on the same day. Students were provided the same amount of time to complete exams regardless of which section (in-residence or distance learning) they were enrolled in. However, distance learning students chose their desired start time individually since there were students in different time zones and with different work schedules. For these students, the professor emailed the exam to the individual at their desired start time. To enforce the allotted time limit, students were required to email their exam responses back when their time was up and were told that one point would be deducted for every five minutes that the students were late. However, all distance learning students abided by the time limit and no deductions had to be enforced.
   a. To make this set-up as unbiased as possible, since the distance-learning students were not proctored during their exams, the exams were

completely open book and open notes. Tests and problems were designed and crafted to discourage memorization or cheating via interactive resources or software applications such as calculators, MATLAB, or Python, which was prohibited.

3. *Access to pre-recorded lectures.* Lectures were previously recorded and made available to all students to access at their leisure and rewatch at any time. For the in-person section, students were encouraged to watch specific pre-recorded lectures before each class. If this occurred, then the allotted class time could then be used for reviewing information, discussing difficult concepts, and answering student questions. For the distance learning section, students were able to choose whichever lectures they preferred: in-person and/or synchronous virtual and/or asynchronous. They were welcome to attend the in-person lecture if they were physically able, watch the asynchronous pre-recorded lectures, or schedule a time to have a synchronous virtual lecture. Any synchronous virtual lectures that were held were recorded and provided to students of both sections.
4. *Student chat page.* An online chat page was provided to all students to discuss lectures, homework, or any questions pertaining to the class. This gave students an autonomous space without the fear, anxiety, or lack of time availability concerns some students may have in asking the professor the same questions. It also boosted study group participation on the practice homework problems.

4.4 Data Demographics

This data set includes six variables: department, military affiliation, in-residence /distance learning (IR/DL), midterm 1 grade, midterm 2 grade, and final exam grade. The response variable is in the form of a continuous grade and a letter grade. There are 46 students (N = 46) total in this data set, three of whom withdrew from the class part way through the quarter. Again, students that withdrew from the class were not included in the analysis and descriptive statistics of the final grade to remain consistent with the approach in Chapter III. However, unlike from the first data set, we know from the instructor when and why these three students withdrew from the class. After receiving their second midterm score and being informed they were on track for a final letter grade lower than a B, they decided to withdraw. It is noteworthy that all three students were in the DL section, although there were several other students in both sections in the same situation who decided to work harder in the remainder of the quarter and managed to turn their grades around. The details of this data set are provided in Table 13.

Table 13: Descriptive Statistics of Response & Variables						
Indicator	Variable	Mean	Median	Min/Max	Std. Dev.	25/75 Quartiles
Y_1	Final Grade - Cont	86.58	88	69/100	7.63	82/92
X_1	Midterm 1	89.42	90	56.67/100	8.44	86.67/96.67
X_2	Midterm 2	77.17	76.67	36.67/100	14.62	70/87.5
X_3	Final Exam	89.24	92.5	55/100	14.42	78.75/92
Indicator	Variable	Response	Freq	Percent	Mean Response	
Y_2	Final Grade - Letter	A	12	26.09%	95.25	
		A-	10	21.74%	89	
		B+	12	26.09%	84.58	
		B	9	19.57%	75	
		W	3	6.52%	---	
X_4	IR/DL	0 (IR)	30	65.22%	87.4	
		1 (DL)	16	34.78%	84.69	
X_5	Military Affiliation	0 (Mil)	41	89.13%	86.82	
		1 (Civ)	5	10.87%	84.8	
X_6	Department	ENG	6	13.04%		
		ENO	1	2.17%		
		ENS	29	63.04%		
		ENY	10	21.74%		

Next, a more in-depth analysis of the response Y_1, Final Grade – Continuous, is conducted. Figure 12 shows the distribution of Y_1 as a histogram and graphed on a normal quantile plot. The response seems to follow the normal distribution fairly well, and a Shapiro-Wilk test confirms this with a p-value of 0.3950, which is larger than the chosen $\alpha = 0.05$. Additionally, the normal quantile plot along with the graph in Figure 13, indicate there are no potential outliers in this data set. For these reasons, a parametric approach will be used for statistical analysis in this data set.

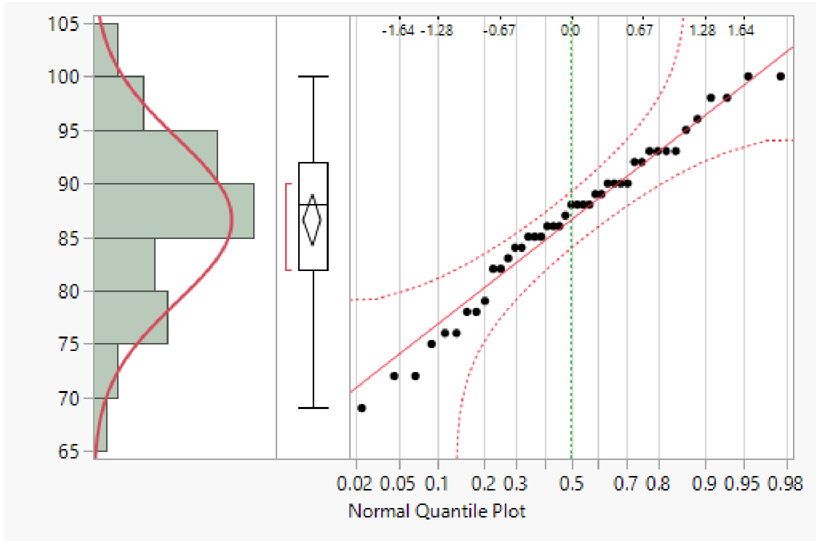


Figure 12: Normal Quantile Plot of Y_1: Final Grade – Continuous II

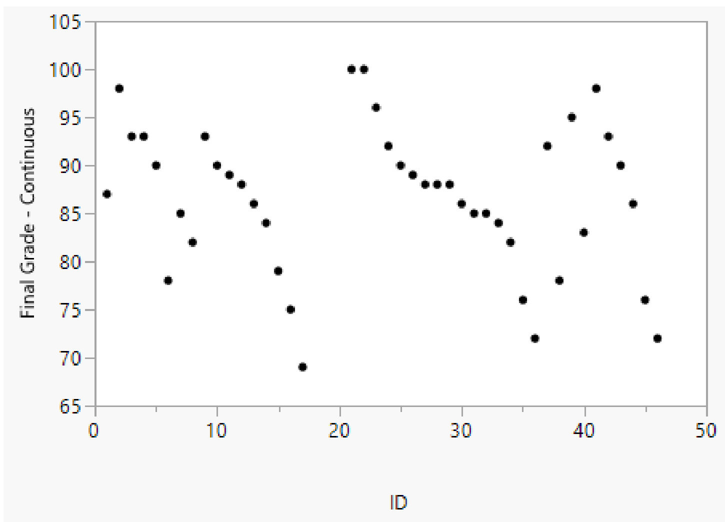


Figure 13: Plot of Y_1: Final Grade – Continuous

4.5 Factor Level Analysis

Before modeling this data set, analysis and comparisons between the response and individual variables are completed. First, consider the relationship between Final Grade – Continuous and IR/DL course status. Figure 14 shows the Final Grade with IR/DL status color coded.

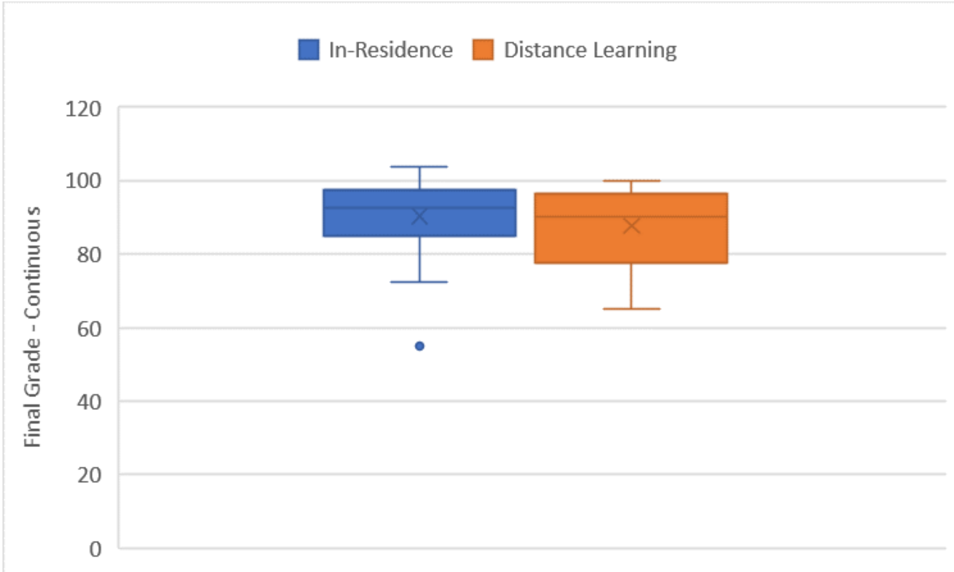


Figure 14: Box Plot of Y_1: Final Grade – Continuous x IR/DL

From Table 13, the Final Grade – Continuous means for the IR/DL sections were 87.4 and 84.69, respectively. However, the independent t-test showed there was no significant difference between the two groups (p-value = 0.298). This mirrors the final conclusions in Chapter III, after modeling was conducted. Shapiro-Wilk tests for Y_1, conditioned on IR and DL sections, confirmed both groups generally followed a normal distribution, with p-values of 0.3804 and 0.7105, respectively. Furthermore, variance testing using the two-sided F-test confirmed there was not a significant difference in variances for the two groups (p-value = 0.9043). The difference in Y_2: Final Grade – Letter when taking the in-residence section compared to the distance learning section is displayed in Figure 15, and also indicates there is not a noteworthy difference when the continuous final grades are converted to letter grades. Pearson Chi-Squared testing solidified the visual interpretation that there is no significant difference in final letter grades based on in-residence and distance-learning status with a p-value of 0.7593.

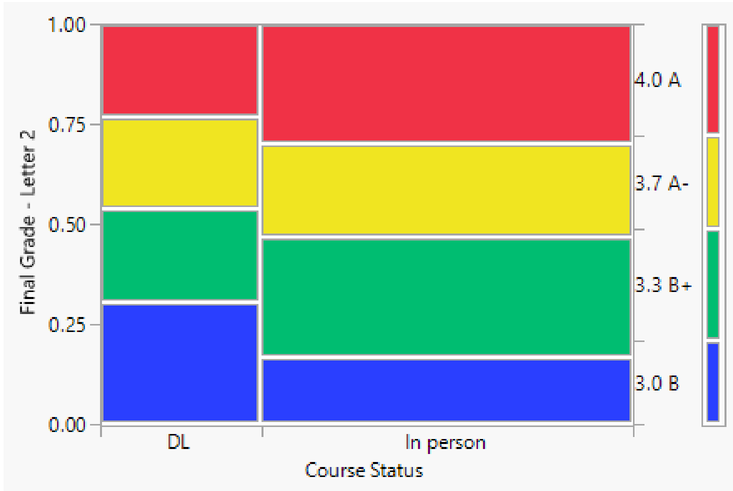


Figure 15: Frequency Plot of Y_2: Final Grade – Letter by Course Status

All of these tests lead to the conclusion that for this particular class and experimental study, there was no significant difference in student performance, as gauged by their final continuous grade, between students enrolled in the in-residence section and the distance learning section.

The same analysis was conducted between Y_1 and X_5, Military Affiliation. Once again, the independent t-test and two-sided F-test infer there is no significant statistical difference in final grades whether a student was military or civilian (p-values of 0.4778 and 0.446, respectively). This, again, matched the conclusions from the first data set in Chapter III regarding final continuous grades and military affiliation.

Lastly, the relationship between Y_1 and X_6, Department, was evaluated. Since there were four groups in X_6, Welch's ANOVA (Analysis of Variance) and student's t tests were conducted. Similar to the previous two results, the results showed there was no significant statistical difference for final grades in variances (O'Brien test p-value = 0.7394) nor means (Welch's ANOVA test p-value = 0.5108), regardless of which department the student belonged to. This conclusion also matched the results from Chapter III, where there were no significant differences in the variances between departments in the initial testing.

Once again, a summary of the tests and the results annotated in this section are displayed in Table 14. The next section discusses the modeling of this data set.

Table 14: Summary of Statistical Tests & Results II

Response x Variable	Type of Test	Test	P-value	Result
Final Grade x X_4: IR/DL	Parametric	F-test 2-sided	0.9043	Equal variances
	Parametric	Independent t-test	0.2980	Equal means
Final Grade x X_5: Military Affiliation	Parametric	F-test 2-sided	0.4460	Equal variances
	Parametric	Independent t-test	0.4778	Equal means
Final Grade x X_6: Department	Parametric	O'Brien	0.7394	Equal variances
	Parametric	Welch's ANOVA	0.5108	Equal means

4.6 Modeling

In considering the previous factor level analysis, one might expect modeling for this data set, with the given variables, to be futile. Indeed, the prediction power of any model created was inadequate in all factors provided. Initial model building indicated there were linear dependencies between X_6: Department and both X_4: IR/DL and X_5: Military Affiliation, so both of those interactions were removed from the model. The full model is shown below:

$$Y_i = \beta_0 + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_{45} X_{4i} X_{5i} + \varepsilon_i$$

Where: Y_i = Final Grade – Continuous, X_4= IR/DL Status, X_5= Military Affiliation, and X_6= Department.

However, attempts to create a reduced model with sequential tests using the chosen α = 0.05 to determine which variables were statistically significant, were not successful since all the p-values for the variables were above the alpha-level. For this reason, further analysis employed the full model. The residuals in this full model did follow a normal distribution and a Shapiro-Wilk test produced a p-value of 0.1118. However, issues were seen with independence of the residuals. Similar to the model building in Chapter III, a Cochrane Orcutt procedure was employed to fix the autocorrelation where $\hat{\rho} = 0.3483997$. The response is transformed and the new response and variables are then refit in the model. However, upon completion of this procedure, an issue with the normality of the residuals is discovered. For this reason, a Box-Cox transformation (also seen in Chapter III) with $\lambda = -0.017$ is conducted. The residuals in this transformed space are seen in Figure 16 and a Shapiro-Wilk test confirms the normality assumption for this model (p-value = 0.0856).

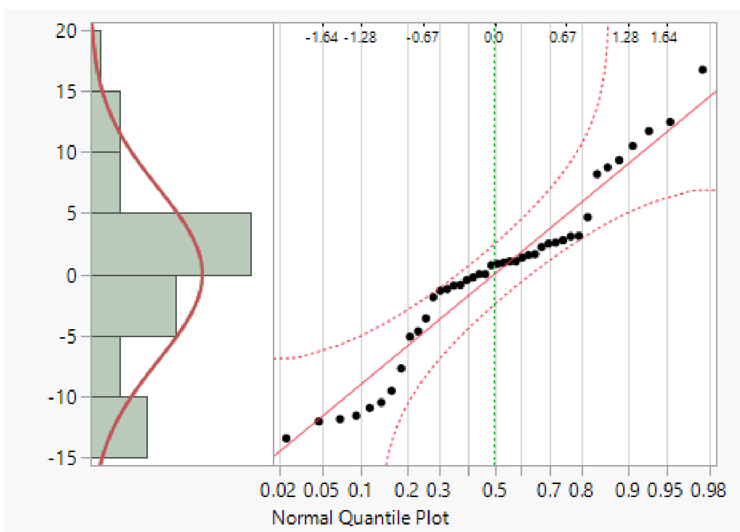


Figure 16: Normal Quantile Plot of Residuals II

The remainder of the model assumptions are rechecked, and no issues are found. The Durbin-Watson test for autocorrelation produces a p-value of 0.1543, confirming there is no serial correlation in the full model. A plot of the residuals against the predicted values for Y, found in Figure 17, confirms this model has constant variance. This plot also confirms non-linearity is not a significant concern.

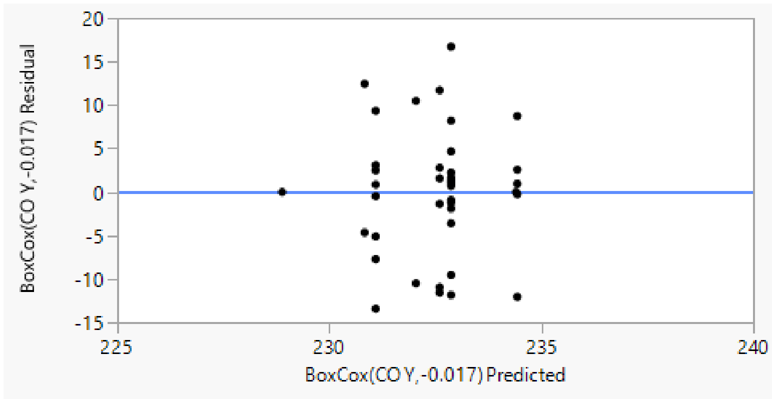


Figure 17: Plot of Y'_t Predicted x Y'_t Residual II

Finally, along with the aforementioned figures, the studentized residuals and the Cook's Distance estimate confirm there are no observations which may be deemed as outliers. Although these conclusions validate the assumptions of linear regression, consider the Parameter Estimates for this model in Table 15, with all p-values well above 0.05.

Table 15: Parameter Estimates II

Term	Estimate	Standard Error	t Ratio	Prob > |t|
Intercept	232.9951	2.558993	91.05	< .0001
Department [ENG]	-0.1393	3.919616	-0.04	0.9719
Department [ENO]	3.812217	8.215465	0.46	0.6455
Department [ENS]	-1.70448	3.028106	-0.56	0.5771
Mil Affil [Civilian]	-0.69161	2.479064	-0.28	0.7819
IR/DL [DL]	0.411959	2.54305	0.16	0.8722
Mil Affil [Civilian]*IR/DL [DL]	1.298886	2.586122	0.5	0.6186

Furthermore, the Summary of Fit and Analysis of Variance details seen in Figures 18 and 19, respectively, indicate the lack of fit and insignificance of this model as a whole. Specifically, the R-squared value of 0.0291 with an adjusted R-squared of -0.1373 and Prob > F (p-value) = 0.9819 demonstrate this model is not useful in predicting students' final grades. There are a few reasons that may have led to modeling being unsuccessful.

Summary of Fit	
RSquare	0.029123
RSquare Adj	-0.13731
Root Mean Square Error	7.637044
Mean of Response	232.4432
Observations (or Sum Wgts)	42

Figure 18: Summary of Fit II

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	6	61.2334	10.2056	0.1750
Error	35	2041.3553	58.3244	**Prob > F**
C. Total	41	2102.5886		0.9819

Figure 19: Analysis of Variance II

First, the sample size of 46 observations is relatively small for the modeling process. This is particularly apparent since some of the groups in the variables have frequencies of ten or less. While the required minimum sample size for each group when conducting comparisons varies between studies and between statisticians, it is common knowledge that the larger the sample size, the more accurate the tests and analysis.

Also, it is likely there were other variables, potentially those that were included in the first data set from Chapter III, that would have been more influential on a student's final grade. The variables that captured previous academic aptitude – undergraduate GPA, GRE Quantitative scores, and the relationship between GPA and GRE Verbal scores – proved to be highly influential on a student's final course grade in Chapter III. Additional variables such as current graduate GPA or math refresher scores may also have been useful in developing a model.

Lastly, this class was rigorously and methodically constructed such that there would be as little differences and biases between the in-residence and distance learning sections as possible. The experimental nature of this study hoped to confirm that when potential biases are controlled, there should not be any significant difference between learning environments – even when also considering the student's department and military affiliation.

4.7 Conclusions

Experimental studies in education are less common than observational studies, as evidenced in the studies covered in Chapter II. Researchers and professors are less willing to manipulate the environment when student performance is the main analysis. For this reason, the study covered in this chapter is a novel addition to the field, with a precise and unbiased manipulation of a course setting that has shown learning environments do not impact student performance when the key elements discussed above are incorporated into a class. This is in consensus with the results and conclusions of Chapter III, where testing with additional variables and modeling proved distance learning students performed similarly to in-residence students when considering final continuous grades. The following chapter will review the conclusions from the whole study and present the implications of those conclusions.

V. Conclusions and Recommendations

5.1 Chapter Overview

This last chapter compiles the conclusions from both data sets and the research conducted from Chapters III and IV, readdressing the research investigative questions presented in Chapter I. Assumptions and limitations of this study are also discussed, along with the implications and significance of this research effort. Finally, the chapter concludes with recommendations for future research.

5.2 Conclusions

The research investigative questions initially presented in Chapter I were the driving force behind the focus of this research. For this reason, they are revisited here and the results and conclusions from both data sets are summarized.

> **1.** How is student performance affected by traditional in-person instruction versus online instruction?

The analysis from Chapters III and IV had different results with respect to this question. Indeed, when utilizing the full data set from Chapter III of 186 observations, a direct comparison between IR/DL and final grade (i.e., when no other variable was accounted for) not only indicated that there was a significant statistical difference in the means between the two groups, but specifically that the final continuous grade mean for distance-learning students was higher than that of in-residence students. Conversely, this same comparison completed on the data set from Chapter IV indicated there was no significant statistical difference in the means between the two groups. However, the differences between the two data sets are noteworthy: the first data set had many forms of bias (most of which were captured in the form of the numerous variables also included), while the second data set was meticulously designed to eliminate as many forms of bias as possible. Additionally, there was a significant difference in the number of observations of each data set, with the first having 186 and the second having only 46. This disparity in the sample size could impact the conclusions from the tests performed, specifically that those tests from the second data set of 46 observations may not be as robust and the margin of error could then be increased.

Ultimately, the conclusion is made that student performance can be affected by traditional in-person instruction versus online or distance learning instruction. However, when the course sections are closely mirrored and taught by the same professor, student performance may not be impacted significantly or at all. The following research question's conclusions consider what happens when those factors and biases are controlled and model assumptions are met.

2. How does controlling for other variables and biases affect the impact instructional technology and learning environment have on student performance?

Since a straight comparison, without additional factors, between in-residence and distance learning students has been considered, now it is imperative to make conclusions about what happens when those potential sources of biases are accounted for. Many of the predictor variables in Chapter III were found to have a significant impact on student performance, as judged by the student's final continuous grade. In fact, it was seen that when just professor is factored in, the difference in students' final continuous grades in relation to in-residence versus distance learning becomes less significant. This implies that while the learning environment could affect student performance, it may be secondary to a professor's teaching styles and preferences. However, because the examinations were not standardized across all the quarters (indicating some professors may create harder questions or grade answers more strictly), this does not translate to a professor's performance as an educator. When additional demographics are considered, such as the student's previous education variables used in Chapter III, undergraduate GPA and GRE scores, learning environment becomes even less substantial in determining a student's final grade. The question of how the increased use of technology over time and as technology improves affect student performance is broached in section 5.5.1.

3. How does AFIT differ from other graduate schools, and how do those differences affect the first two questions?

The beginning of Chapter III discussed some of the differences and similarities between AFIT and other graduate schools. In considering the variables that were found to be significant factors in a student's performance, it is likely the differences between AFIT and other institutions could have a profound impact on student performance. The studies found in the literature review had mixed results and conclusions, but comparisons can be made between the two data sets in this study and those studies from Chapter II.

First, there were two studies from graduate schools: Hoffman & Elmi, 2021 and Stephenson, 2001. Both studies came to the same conclusion: that there was no significant statistical difference between course environments. This conclusion mirrors those from the first data set after modeling and the second data set with direct comparisons. This may imply that the demographic and institutional differences that make AFIT-EN unique may not have as much of an impact on student performance and IT as initially postulated by the author. However, the establishment of those two studies was notably different than this research. Both studies had only one instructor and collected data over the course of five years, with much larger sample sizes of 412 and 1,759 students. Nonetheless, comparisons are provided between this research on AFIT-EN and the studies with other degree program levels (undergraduate, high school, and combinations) that more closely align with each data set in this research on other aspects.

The studies found that were most closely related to the first data set were those conducted on one course, with multiple professors over multiple semesters. There were three studies referenced in Chapter II that fell under these categories: Farmus and others, 2020, Hake, 1998 and Winquist & Carlson, 2014, all of which concluded the courses which utilized instructional technology performed better. This is in line with the initial comparison completed in Chapter III between only final grade and IR/DL. However, unlike the remainder of the analysis, modeling, and conclusions from Chapter III, none of those studies made further analyses with student, instructor, or class demographics. Therefore, it cannot be ascertained if differences in AFIT caused the discrepancy between the results presented here and in those studies, or whether similar results would have been obtained in the three previous studies had they accounted for variables such as instructor and student demographics.

Similarly, the studies from Chapter II that most resembled the second data set were those conducted on one course, with one professor over one semester in which exam scores and final grades were used to conduct the analysis. There were two studies that had these characteristics, which had differing conclusions: Dutton & Dutton, 2005 and Gundlach and others, 2015. Both studies were carefully designed such that both the distance learning and the in-residence sections were as similar as possible. However, both studies had much larger sample sizes (200-500 students total) than the 46 students from the data set from Chapter II. Even with the variation in sample sizes, since one study concluded there was no significant statistical difference in course sections and the other concluded courses with IT performed better, they are still on par with the conclusions made in the second data set of this study.

While similarities can be made between this study and the studies in the literature review, the few studies found that were conducted on graduate schools, were not a direct match for comparing. For this reason, the remainder of the conclusion for this final research question is left to future research to conduct and analyze similar studies at other graduate schools that more closely resemble AFIT.

5.3 Assumptions and Limitations

Both sets of data used in this research came from AFIT-EN, which has unique characteristics that are assumed to have an impact on student performance, as discussed in Chapter II. This assumption also implies a limitation in any conclusions that are made. Therefore, any conclusions stated in this research should only be applied to similar situations and institutions. This could comprise of, but is not limited to, applying these conclusions to science, technology, engineering, and mathematics (STEM) courses at a university with students of comparable education backgrounds as those found at AFIT and within this study.

Moreover, there may be a few forms of bias in this research that could not be accounted for within the analysis. Instructor preference for a specific way of teaching, student preference for a specific way of learning, and the inability to randomize students into specific course sections are potential biases discussed in the referenced studies in Chapter II.

Additionally, there was no variable that captured time in the first data set in Chapter III. Not only is this a possible cause of the autocorrelation found while model building, but this also meant there was an assumption within the study that time was not a concern. However, as discussed in the beginning of Chapter I, technology continues to improve over time and students and instructors alike have steadily more experience with utilizing technology as a learning tool. An ideal data set to address technology expansion would not only have a variable to capture time, but also a variable to capture the quantity and quality of the technology that was utilized in each class section.

5.4 Implications

This research has the potential to be applied to future classes to enable improved student performance, whether for in-residence courses or distance learning courses. For example, when designing a class that is prescribed to have both types of course sections (or any combination of learning environments), the main elements from Chapter IV could be included. This may help to ensure all students, regardless of enrollment in a specific section or educational background have the same access to information and tools and have the same opportunities as their counterparts in another course section.

In addition, this research can be employed by instructors to assist them in understanding where and when students may struggle. For example, finding an ideal time within the quarter to talk to students about their grades and potentially needing to drop the course or work harder is not easy or simple. During the quarter the second data set was collected in, the instructor informed students who were at risk of receiving a grade below a B after the second midterm. This gave students sufficient time to make a decision about their enrollment and dedication to completing the class with a satisfactory grade, and ultimately resulted in a few students withdrawing and all other student achieving a B or higher.

Furthermore, students may peruse this study and apply the information and conclusions to their own experiences and future education. When given options, choosing which course section to enroll in with no background on how the differences might impact yourself can be a daunting task. While the main conclusion of this study was that there is no difference in student performance in respect to classroom environment, this was somewhat contingent on different factors in both data sets. Whether student educational background or course characteristics, speaking with the professor(s) of the course is a fine place to start. Finding the right course section for themselves can set students up for success, both in that particular class and in long term educational goals.

5.5 Recommendations for Future Research

Results of this study may also be utilized to support future research into other instructors, courses, departments, or institutions. While there has already been a myriad of studies performed to assess student performance and learning environments, there is always more research to be completed that can further help the general understanding of this topic and ensure student performance is prioritized.

5.5.1 Additional Predictor Variables

First, as has already been alluded to in other sections of this research, including additional potential predictor variables could be highly beneficial to uncovering more concrete conclusions. The first half of this study was able to include the previous education markers of undergraduate GPA and GRE Quantitative and Verbal scores. Other previous education markers to consider for inclusion may comprise of: math refresher course grades and current graduate GPA if it isn't the student's first semester in the graduate school. A variable to capture knowledge retention rates, potentially from a post-test conducted some time after course completion, may also provide valuable insight.

Another potential variable is the number of distance learning courses the student has previously taken, as this experience may also significantly impact a student's performance in a distance learning course section. While degree type (non-degree, certificate, masters, or PhD) was collected in this study in the form of class totals, it wasn't attached to specific students due to the non-identifiability requirement. Accounting for the type of degree a student is pursuing may help to further explain variability in the model building process. Furthermore, current student course workload was included in some studies in Chapter II and found to have an impact on student performance. This type of variable, and potentially even including external work requirements as well, could be beneficial to understanding how mental and time responsibilities impact student performance, dependent on learning environments.

Future research could also consider factors related to instructors. Collecting demographics from professors may include, but is not limited to: how many distance learning courses or courses in general the professor has taught at the time of that course, total years of teaching, how many semesters they taught the specific course under evaluation, what instruction technology they have used, what/how many teaching courses they have taken, and any preferences for teaching style or learning environment. Until further research is completed with these potential variables, it is speculation as to which ones, if any, would be significant in the outcome of student performance.

Questionnaires for both instructors and students could also prove essential in an analysis. While this study was focused on student performance, student and teacher fulfillment, gratification, and motivation (intrinsic or extrinsic) are also important insights into how technology use impacts education. A student may perform very well in a distance learning class but receive no enjoyment or fulfillment from taking the class, which could impact future classes the student decided to take. Potential questions for students and teachers on questionnaires may include:

- Did you enjoy the technology used in this course?
- How satisfied are you with the application of the technology used?
- I/my professor were well prepared to utilize the technology required in this class.
- The software utilized was easy to acquire and navigate.

- I feel confident in utilizing technology for teaching/learning.

To include the answers in a statistical analysis, answers could be presented on a Likert scale. Providing five (or seven) possible answers to a statement or question, the Likert scale allows the student or professor to: "indicate their positive-to-negative strength of agreement or strength of feeling regarding the question or statement" (Mcleod, 2019). This is commonly seen as a strongly disagree (1) to strongly agree (5) response, although additional Likert scale examples are provided in Figure 20. This type of questionnaire and response provides the researcher with a tangible variable in the form of a one to five (or one to seven) numerical, ordinal response to include in their study and analysis.

Response Set	1	2	3	4	5
Frequency	Never	Rarely	Sometimes	Often	Always
Quality	Very poor	Poor	Fair	Good	Excellent
Intensity	None	Very mild	Mild	Moderate	Severe
Agreement	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Approval	Strongly disapprove	Disapprove	Neutral	Approve	Strongly approve
Awareness	Not at all aware	Slightly aware	Moderately aware	Very aware	Extremely aware
Importance	Not at all important	Slightly important	Moderately important	Very important	Extremely important
Familiarity	Not at all familiar	Slightly familiar	Moderately familiar	Very familiar	Extremely familiar
Satisfaction	Not at all satisfied	Slightly satisfied	Moderately satisfied	Very satisfied	Completely satisfied
Performance	Far below standards	Below standards	Meets standards	Above standards	Far above standards

Figure 20: Likert Scale Examples

Section 5.3 referred to the limitation of not including a variable to capture time, which also ties into recommendations for future research. This could be in the form of the semester and/or year the course occurred. It also might be in the form of how user-friendly the technology used in the courses was. If the different course sections employed the same or similar software applications, it could also be in the form of which version of the software was utilized. This could be a serious indicator of whether technology is becoming more or less useful in the classroom.

5.5.2 Expanding the Subset

While exploring additional potential predictor variables for inclusion in analysis and model building is undoubtedly vital to future research, there are other ways to expand the conclusions in this field of study. For example, the study conducted in this research focused on a subset of the academic world: one course from one department of a specific graduate institution. While this is how many studies are conducted, expanding this subset to include other instructors, courses, or departments may produce more robust conclusions. More singular instructor comparisons could help professors find their weaknesses and strengths, and find ways to capitalize on their strengths and improve their weaknesses.

Department-wide studies in which all courses taught by that department are included would indicate how that department compares to others. This would also give insight into how other courses, such as engineering, physics, and chemistry courses, compare with mathematics courses when it comes to learning environment.

AFIT-EN-wide studies would show how AFIT as a whole institution handles instructional technology and distance learning courses and how it impacts student performance. This may help determine whether AFIT is on the same standing as other research institutions when it comes to supporting students and professors with utilizing technology in the classrooms.

5.5.3 Improving IT at AFIT

If further research indicates students and professors alike believe improvements could be made or that AFIT is not up to standard when it comes to handling different learning environments, the question should then be asked: how can AFIT make teaching with technology better and easier? The interviews with professors from Section 2.8 alluded to this by making some suggestions of capabilities they would like to have which are not currently available to them. These included: more technology support for professors and students, rooms and equipment for recording lectures in, better technology (to include software and hardware), and faster and more reliable Wi-Fi. These additions and improvements would benefit classes when incorporating technology in academia at AFIT. Future research could expound on this to determine how student performance changes with better technology and more prepared instructors.

5.6 Summary

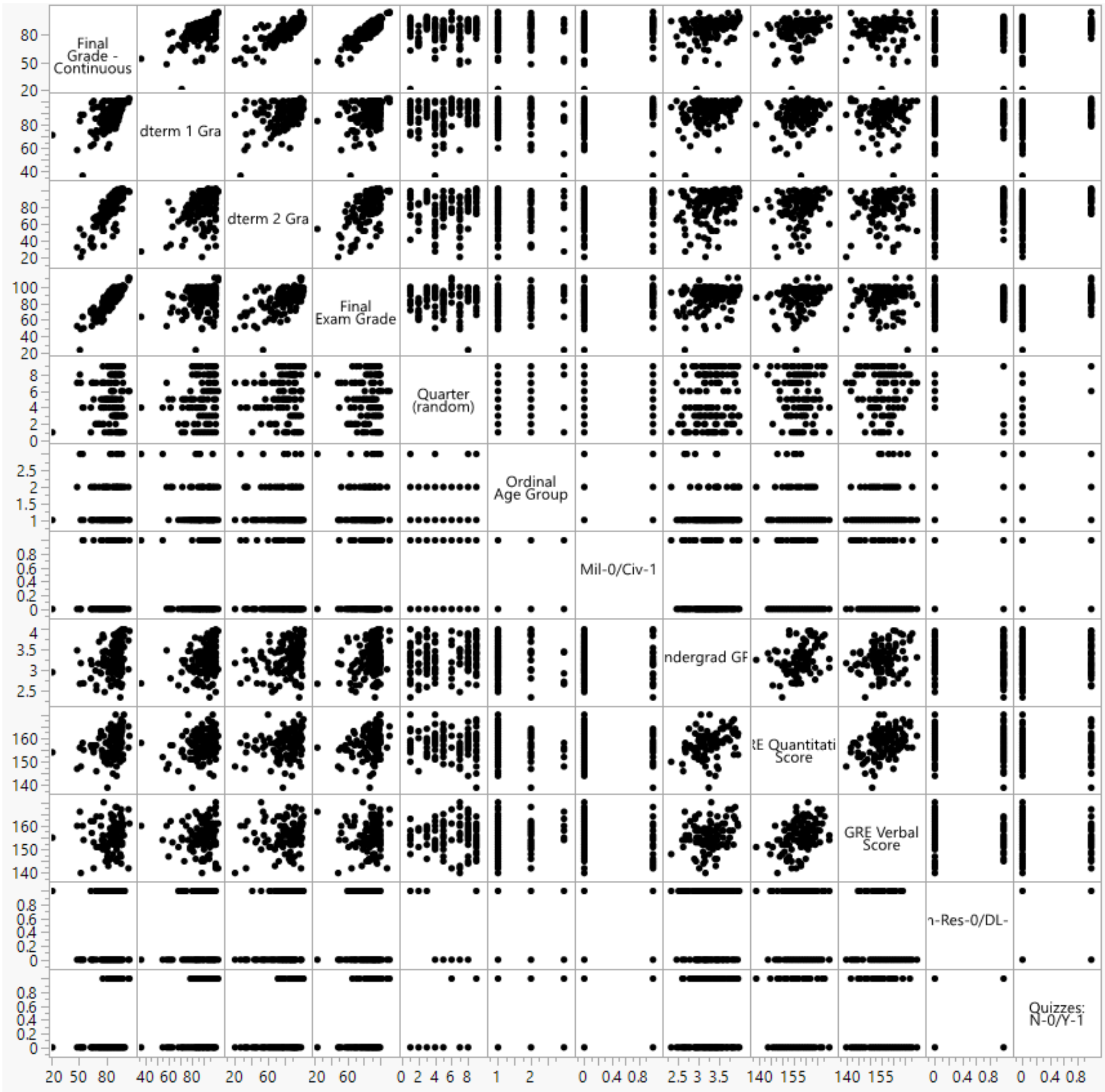
Technology has seen a significant change in the last few decades, but exploring how that growth in technology effects student performance is a vital aspect of education, specifically when utilizing technology in education. By studying the impacts and consequences of instructional technology, education as a whole can be improved. Whether more technology is used or less, whether specific software or hardware is used for specific institutions, the impact that technology has on student performance is better understood. In turn, student satisfaction, future education, and success can also be improved.

Appendix A: Summary of Literature Review Studies

Study Author	Course Level	Course	# Professors	# Semesters	Scores Included	Demographic Data Included*
No Significant Statistical Difference						
Hoffman & Elmi, 2021	Graduate	Introductory Biostatistics	1	20 (5 years)	Math Refresher, Quizzes, Final Exam, Lab Participation	Sex, Race/Ethnicity, Age, Student Hours (full/part-time), Quantitative GRE, Undergraduate GPA
Stephenson, 2001	Graduate	Applied Statistics for Industry	1	10 (5 years)	Grade Point Average, Letter Grade	Student Hours (full/part-time)
Merisotis & Phipps, 1999	Various	Various	40	40 studies	Various	N/A
Gundlach et al., 2015	Undergraduate	Statistical Literacy	1	1	Exams, Homework, and Project Grades	Student Classification Level (freshman-senior), Age, Sex, Race, Domestic/International Status
IT Courses Perform Better						
Farmus et al., 2020	Undergraduate	Introductory Statistics	11	11 studies	Final Grade, Final Exam	None
Strayer, 2012	Undergraduate	Introductory Statistics	1	1	**Quantifiable scores were not utilized	None
Wilson, 2013	Undergraduate	Introductory Statistics	1	1 (4 sections)	Exam Grades, Final Grade	None
Winquist & Carlson, 2014	Undergraduate	Introductory Statistics	2	11 years	Psychology Area Concentration Achievement Test (ACAT) with 10 scales	None
Deslauriers et al., 2011	Undergraduate	Engineering Calculus	2	1	Exams	None
Hake, 1998	Undergraduate & High School	Introductory Physics	Various	48 data sets	Force Concept Inventory (FCI) and Mechanics Baseline (MB) tests	Institution Level
Dutton & Dutton, 2005	Undergraduate	Economics & Business Statistics	1	1	Final Exam, Weighted Average of Tests & Projects	GPA, Course Hours, Major, Age, Gender, Computer Experience, Work/Child Care Commitments, Commuting Distance
Lazari, 2018	Undergraduate	Algebra	1	2	Final Exam	Retention Rates
IT Courses Perform Worse						
Tanyel & Griffin, 2014	Undergraduate	Various	38	15	Final Grade	Major, Prior GPA, Age
Bettinger et al., 2015	Undergraduate	Various	Various	5 years	Final Grade	Age, Sex, Student Hours, Prior GPA, Home Campus
Kofoed et al., 2021	Undergraduate	Principles of Economics	12	1	Final Grade	Sex, Race/Ethnicity, Athlete Status, Prior Enlisted, CEER, Class Time/Day, Instructor Effect

*: Not all of the demographic data listed was utilized in analyzing performance; some was included to reflect the sample size.

Appendix B: Scatterplot Matrix



Appendix C: IRB Approvals

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY (AETC)

HRPP Exempt Determination Form

For AFIT HRPP Use Only	
Protocol Number:	
Protocol Title:	

EDO Determination

Does this submission meet an Exempt Criteria? *Select the appropriate exemption category. Categories are defined in Exemption Request Package and on Page 2 of this form.*

☑ Yes	Which exempt category applies?	32 CFR 219.104 (d) (1) (i)
	Is a limited IRB Review required to determine adequate provisions are in place to protect the privacy of subjects and maintain confidentiality of data?	☐ Yes ☑ No
	If a limited IRB review is required, IRB Member determined that *either*: ☐ Sufficient measures were taken to protect privacy and confidentiality. - OR - ☐ Insufficient measures were taken to protect privacy and confidentiality.	
☐ No	☐ The human subject research does not meet any exempt criteria. Referred to AFRL IRB Chair for IRB review.	
	- OR -	
	☐ The research uses an In Vitro diagnostic device with specimens that are NOT individually identifiable. Referred to AFRL IRB Chair to determine compliance with applicable FDA regulations.	

AFIT EDO / IRB Member Submission Analysis

EDO Reviewer Comments

The research seems to be exempt under exempt category 1. This research seems to involve an established accepted educational setting with normal educational practices that will bnot impact negatively on the students' opportunity to learn required educational content or the assessment of educators. Demographic information that is collected is not outside of normally collected data in a classroom setting, and at any rate, privacy will be insured. The research is concerned with differences in delivery methods, both of which are currently, and in the past, been used at AFIT. Comparisons are to be made to ascertain any differences in instructional techniques.

AFIT EDO Signature

CUNNINGHAM.WILLIAM.A.III.1230804660 Digitally signed by CUNNINGHAM.WILLIAM.A.III.1230804660 Date: 2022.08.03 09:56:00 -04'00'	Click or tap to enter a date.
Exempt Determination Official	Date 3 Aug 2022

Note: To sign this form electronically, please save it as a PDF and follow these instructions.

20 June 2019 1

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY (AETC)

HRPP Exempt Determination Form

For AFIT HRPP Use Only	
Protocol Number:	
Protocol Title:	

EDO Determination		
Does this submission meet an Exempt Criteria? *Select the appropriate exemption category. Categories are defined in Exemption Request Package and on Page 2 of this form.*		
☑ Yes	Which exempt category applies?	**32 CFR 219.104 (d) (3)** (i)
	Is a limited IRB Review required to determine adequate provisions are in place to protect the privacy of subjects and maintain confidentiality of data?	☐ Yes ☑ No
	If a limited IRB review is required, IRB Member determined that *either*: ☐ Sufficient measures were taken to protect privacy and confidentiality. - OR - ☐ Insufficient measures were taken to protect privacy and confidentiality.	
☐ No	☐ The human subject research does not meet any exempt criteria. Referred to AFRL IRB Chair for IRB review.	
	- OR -	
	☐ The research uses an In Vitro diagnostic device with specimens that are NOT individually identifiable. Referred to AFRL IRB Chair to determine compliance with applicable FDA regulations.	

AFIT EDO / IRB Member Submission Analysis
EDO Reviewer Comments
The research seems to be exempt under exempt category 3. Any disclosure of the human subjects' responses outside the research would not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, educational advancement, or reputation, and the researcher is taking steps to safeguard the information.

AFIT EDO Signature	
CUNNINGHAM.WILLIAM.A.III.1230804660 Digitally signed by CUNNINGHAM.WILLIAM.A.III.1230804660 Date: 2022.09.26 11:46:45 -04'00'	Click or tap to enter a date.
Exempt Determination Official	Date 23 Sep 2022
Note: To sign this form electronically, please save it as a PDF and follow these instructions.	

20 June 2019 1

Appendix D: Abbreviated Consent Form

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY (AETC)

Abbreviated Consent
for
Student Performance in Traditional vs. Online Sections of an
Introductory Graduate Mathematics Course

You are being asked to participate in a research study.

Key study information you should know:

- The purpose of the study is to determine if student performance is affected by traditional in-person instruction versus online instruction and if so, how. If you choose to participate, you will be asked to answer a handful of interview questions revolving around your experiences and opinions on instructional technology and online learning. This will take approximately 30 minutes.

- Risks or discomforts from this research are expected to be negligible since no personal identifiable information will be available to non-researchers.

- The study will have no direct benefits to the subjects.

- Taking part in this research project is voluntary. You can discontinue participation at any time without penalty or loss.

Please take time to read this entire document and ask questions before deciding whether to take part in this research project.

If you participate in this research, the researcher (on behalf of the Principal Investigator) will contact you to set up an interview time. You may choose to answer the interview questions at your own leisure, via email. If you choose to answer via email, possible standard follow-on questions to responses may be sent. If you choose to be interviewed orally, those follow-on questions will be asked at that time.

The researchers will take the following precautions to maintain the confidentiality of your data: No participant identifiers linked to you will be collected nor included in any publications.

The data may be accessed by the Department of Defense for auditing purposes.

If you have questions regarding the study, contact the Principal Investigator: Major Victoria Sieck, 509-768-4853 or Victoria.sieck@afit.edu. If you have questions regarding your rights as a research subject, contact the AFIT HRPP: 937-255-3636 x4543 or humansubjects@afit.edu.

Appendix E: Instructor Questionnaire

Questionnaire for Research Study:
Student Performance in Traditional vs Online Sections of an
Introductory Graduate Mathematics Course

1. What instructional technology (IT) have you used?
   Please be as specific as possible (e.g., MS Teams, SMART Boards, chat rooms, online textbooks, pre-recorded video lectures, collaborative software such as Google Docs, CBTs, etc.)

2. How many classes have you taught using IT? How many classes have you taught that have been fully distance learning?

3. What could be/is different about using IT at AFIT vs elsewhere?

4. What difficulties have you come across while teaching with IT?

5. What advantages have you come across while teaching with IT?

6. What is your preferred method of teaching? Why?

7. Do you have any recommendations for teachers who are newly integrating IT into their curriculum?

8. What IT capabilities would you like to have that are not currently available to you?

9. Any other comments you would like to make?

Bibliography

5 Different Types of Educational Technology | TL Dev Tech. (2022, April 5). https://www.tldevtech.com/different-types-of-educational-technology/

About the Air Force Institute of Technology. (n.d.). AFIT.Edu. Retrieved February 1, 2023, from https://www.afit.edu/ABOUT/

AFIT / Office of Admissions / Air and Space Force. (n.d.). EN Office of Admissions. Retrieved January 6, 2023, from https://www.afit.edu/ADMISSIONS/page.cfm?page=472&tabname=Tab3A

AFIT ENE Admissions. (2022, October 31). *Request for Information—MS Student Research* [Personal communication].

Ahmadigol, J. (2016). *New Definition of Educational Technology*. 3.

Association for Educational Communications and Technology. (n.d.). Retrieved July 27, 2022, from https://aect.org/news_manager.php?page=17578

Ballard, C. (2019, July 5). *An Ode To R-Squared*. Towards Data Science. https://towardsdatascience.com/an-ode-to-r-squared-804d8d0ed22c

Bates, A. W. (Tony). (2005). *Technology, E-learning and Distance Education*. Psychology Press.

Bedwell, W. L., & Salas, E. (2010). Computer-based training: Capitalizing on lessons learned. *International Journal of Training and Development*, *14*(3), 239–249. https://doi.org/10.1111/j.1468-2419.2010.00355.x

Bettinger, E., Fox, L., Loeb, S., & Taylor, E. (2015). Changing Distributions: How Online College Classes Alter Student and Professor Performance (CEPA Working Paper No.15-10). *Standford Center for Education Policy*, 33.

Boyd, D. J., Grossman, P. L., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher Preparation and Student Achievement. *Educational Evaluation and Policy Analysis*, *31*(4), 416–440. https://doi.org/10.3102/0162373709353129

Brame, C. J. (2013). *Flipping the Classroom*. Vanderbilt University. https://cft.vanderbilt.edu/guides-sub-pages/flipping-the-classroom/

Burke, S. (1998). Missing Values, Outliers, Robust Statistics & Non-parametric Methods. *Scientific Data Management*, *1*, 32–38.

Carlton, G. (2022, September 29). *What Is A Student-to-Professor Ratio In College?* Affordable Colleges. https://www.affordablecollegesonline.org/college-resource-center/student-faculty-ratio-finding-the-right-class-size/

Casella, G., & Berger, R. L. (2002). *Statistical Inference* (Second). Cengage Learning.

Chen, N., Ko, H., Kinshuk *, & Lin, T. (2005). A model for synchronous learning using the Internet. *Innovations in Education and Teaching International*, *42*(2), 181–194. https://doi.org/10.1080/14703290500062599

Cheong, C., Coldwell-Neilson, J., MacCallum, K., Luo, T., & Scime, A. (2021). *COVID-19 and Education: Learning and Teaching in a Pandemic-Constrained Environment*. Informing Science.

Cristea, M.-A. (2020, March 26). How Has Technology Impacted The Medical Field? *Business Review*. https://business-review.eu/business/healthcare/how-has-technology-impacted-the-medical-field-208985

D’Agostino, S. (2022, October 20). *How COVID Spurred Digital Innovation and Empathy*. Inside Higher Ed. https://www.insidehighered.com/news/2022/10/20/how-covid-spurred-online-education-innovation-and-empathy

de Vries, T. J. (2021). The Pandemic That Has Forced Teachers to Go Online. Zooming in on Tips for Online Teaching. *Frontiers in Education*, *6*. https://www.frontiersin.org/articles/10.3389/feduc.2021.647445

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, *113*(3), 554–559. https://doi.org/10.1073/pnas.1517441113

DeRuy, E., & National Journal. (2015, October 28). *At Universities, More Students Are Working Full-Time*. The Atlantic. https://www.theatlantic.com/politics/archive/2015/10/at-universities-more-students-are-working-full-time/433245/

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved Learning in a Large-Enrollment Physics Class. *Science*, *332*(6031), 862–864. https://doi.org/10.1126/science.1201783

Difference Between Educational and Instructional Technology. (n.d.). Retrieved May 3, 2022, from https://ecampus.unza.zm/blog/page/2/, https://ecampus.unza.zm/blog/page/3/, https://ecampus.unza.zm/educational-instructional-technology-difference/

Dogan, U. (2015). Student Engagement, Academic Self-efficacy, and Academic Motivation as Predictors of Academic Performance. *The Anthropologist*, *20*(3), 553–561. https://doi.org/10.1080/09720073.2015.11891759

Dong, Y., & Peng, C.-Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, *2*, 222. https://doi.org/10.1186/2193-1801-2-222

Dutton, J., & Dutton, M. (2005). Characteristics and Performance of Students in an Online Section of Business Statistics. *Journal of Statistics Education*, *13*(3), null. https://doi.org/10.1080/10691898.2005.11910564

Faculty Retention Proves a Major Challenge for Universities. (n.d.). Retrieved January 31, 2023, from https://news.rpi.edu/luwakkey/2994

FAQ: Instructional Technology Program, Watson College of Education at the University of North Carolina Wilmington. (n.d.). Retrieved July 27, 2022, from https://uncw.edu/ed/mit/faq.html

Farmus, L., Cribbie, R. A., & Rotondi, M. A. (2020). The Flipped Classroom in Introductory Statistics: Early Evidence From a Systematic Review and Meta-Analysis. *Journal of Statistics Education*, *28*(3), 316–325. https://doi.org/10.1080/10691898.2020.1834475

Frost, J. (2017a, May 10). How High Does R-squared Need to Be? *Statistics By Jim*. http://statisticsbyjim.com/regression/how-high-r-squared/

Frost, J. (2017b, October 31). Understanding Interaction Effects in Statistics. *Statistics By Jim*. http://statisticsbyjim.com/regression/interaction-effects/

FY2022 AFIT Fact Sheet. (2022). https://www.afit.edu/docs/FY2022_AFIT_Fact_Sheet_Final.pdf

Gagne, R. M. (1987). *Instructional Technology: Foundations* (1st ed.). Routledge.

Grace-Martin, K. (2013, September 18). The Difference Between Crossed and Nested Factors. *The Analysis Factor*. https://www.theanalysisfactor.com/the-difference-between-crossed-and-nested-factors/

Graduate School of Engineering and Management Fall 2021 Fact Book. (2022). https://www.afit.edu/images/pics/file/Fall%202021%20Fact%20Book.pdf

Groff, J., & Mouza, C. (2008). A Framework for Addressing Challenges to Classroom Technology Use. *AACE Review (Formerly AACE Journal)*, *16*(1), 21–46.

Gundlach, E., Richards, K. A. R., Nelson, D., & Levesque-Bristol, C. (2015). A Comparison of Student Attitudes, Statistical Reasoning, Performance, and Perceptions for Web-Augmented Traditional, Fully Online, and Flipped Sections of a Statistical Literacy Class. *Journal of Statistics Education*, *23*(1), null. https://doi.org/10.1080/10691898.2015.11889723

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64–74. https://doi.org/10.1119/1.18809

Hatziapostolou, T., & Paraskakis, I. (2010). Enhancing the Impact of Formative Feedback on Student Learning Through an Online Feedback System. *Electronic Journal of E-Learning*, *8*(2), 111–122.

Hernández González, C. A., & Blackford, B. J. (2022). Engagement as antecedent of academic achievement and the moderating impact of work-family-school inter-role conflict for online graduate students. *The International Journal of Management Education*, *20*(3), 100676. https://doi.org/10.1016/j.ijme.2022.100676

Hockly, N. (2018). Blended Learning. *ELT Journal*, *72*(1), 97–101. https://doi.org/10.1093/elt/ccx058

Hoffman, H. J., & Elmi, A. F. (2021). Comparing Student Performance in a Graduate-Level Introductory Biostatistics Course Using an Online versus a Traditional in-Person Learning Environment. *Journal of Statistics and Data Science Education*, *29*(1), 105–114. https://doi.org/10.1080/10691898.2020.1841592

Hollander, M., & Sethuraman, J. (2015). Nonparametric Statistics: Rank-Based Methods. In *International Encyclopedia of the Social & Behavioral Sciences* (Second, Vol. 16, pp. 891–897). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.42157-4

Hollander, M., Wolfe, D. A., & Chicken, E. (2014). *Nonparametric Statistical Methods* (Third). John Wiley & Sons.

Hoskin, T. (2012). Parametric and Nonparametric: Demystifying the Terms. *Mayo Clinic*, *5*(1), 1–5.

Hu, J. C. (2020, June 24). *Graduate programs drop GRE after online version raises concerns about fairness*. Science.Org. https://www.science.org/content/article/graduate-programs-drop-gre-after-online-version-raises-concerns-about-fairness

Ilhan, G. O., & Oruç, Ş. (2016). Effect of the use of multimedia on students' performance: A case study of social studies class. *Educational Research and Reviews*, *11*(8), 877–882.

Jaffee, D. (1997). Asynchronous Learning: Technology and Pedagogical Strategy in a Distance Learning Course. *Teaching Sociology*, *25*(4), 262–277. https://doi.org/10.2307/1319295

Januszewski, A. (2008). *Educational Technology: A Definition with Commentary*. Routledge.

Karaca, C., & Ocak, M. A. (2017). *Effect of Flipped Learning on Cognitive Load: A Higher Education Research*. 8.

Khanna, C. (2020, December 5). *Observational vs Experimental Study*. Medium: Towards Data Science. https://towardsdatascience.com/observational-vs-experimental-study-543425a3b3c8

Kim, J. H. (2015). *How to Choose the Level of Significance: A Pedagogical Note*. Munich Personal RePEc Archive. https://mpra.ub.uni-muenchen.de/69992/

Kofoed, M. S., Gebhart, L., Gilmore, D., & Moschitto, R. (2021). Zooming to Class?: Experimental Evidence on College Students' Online Learning during COVID-19. *IZA Discussion Papers*, *14356*, 36.

Korucu-Kis, S., & Ozmen, K. S. (n.d.). *S-TECHNO: An instructional design model for redesigning instructional technology courses*. 19.

Kowarski, I. (2021, April). *What the GRE Test Is and How to Prepare*. US News & World Report. https://www.usnews.com/education/best-graduate-schools/articles/what-the-gre-test-is-and-how-to-prepare

Kruse, C. S., & Beane, A. (2018). Health Information Technology Continues to Show Positive Effect on Medical Outcomes: Systematic Review. *Journal of Medical Internet Research*, *20*(2), e8793. https://doi.org/10.2196/jmir.8793

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2004). *Applied Linear Statistical Models* (Fifth). McGraw-Hill/Irwin.

Laal, M., & Ghodsi, S. M. (2011). Benefits of collaborative learning. *Procedia - Social and Behavioral Sciences*, *31*, 491–495. https://doi.org/10.1016/j.sbspro.2011.12.091

Laal, M., & Laal, M. (2011). Collaborative learning: What is it? *Procedia - Social and Behavioral Sciences*, *31*, 491–495. https://doi.org/10.1016/j.sbspro.2011.12.092

Lane, H., Tina, Richard, H., & Richard, T. (2019). *Handbook of Research on Emerging Practices and Methods for K-12 Online and Blended Learning*. IGI Global.

Lazari, A. (2018). College Algebra – Online Section Versus Traditional Section. *Georgia Journal of Science*, *76*, 7.

Lee, B. N., Abdullah, S., & Na Kiu, S. (2016). Supportive Learning: Linear Learning and Collaborative Learning. *The Turkish Online Journal of Educational Technology*, *15*(2), 9.

Li, X., Wang, F., Xu, C., & Zhao, C. (2009). The Postgraduate Specialty Curriculum Arrangement Analysis of Instructional Technology in America. *2009 First International Conference on Information Science and Engineering*, 3512–3515. https://doi.org/10.1109/ICISE.2009.1242

Limantara, N., Kosala, R., Ranti, B., & Supangkat, S. H. (2019). Human and Technology Factors in the Readiness to Use Smart Mobile Devices in Learning Activities. *2019 International Conference on ICT for Smart Society (ICISS)*, *7*, 1–5. https://doi.org/10.1109/ICISS48059.2019.8969827

Lum, H. C., Elliott, L. J., Aqlan, F., & Zhao, R. (2020). Virtual Reality: History, Applications, and Challenges for Human Factors Research. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *64*(1), 1263–1268. https://doi.org/10.1177/1071181320641300

Madley-Dowd, P., Hughes, R., Tilling, K., & Heron, J. (2019). The proportion of missing data should not be used to guide decisions on multiple imputation. *Journal of Clinical Epidemiology*, *110*, 63–73. https://doi.org/10.1016/j.jclinepi.2019.02.016

Maier, M., & Lakens, D. (2022). Justify Your Alpha: A Primer on Two Practical Approaches. *Advances in Methods and Practices in Psychological Science*, *5*(2). https://doi.org/10.1177/25152459221080396

McLeod, S. A. (2019, July 4). *What are Type I and Type II Errors? - Simply Psychology* [Simply Psychology]. https://www.simplypsychology.org/type_I_and_type_II_errors.html

Mcleod, S. A. (2019). Likert Scale. *Simply Psychology*. https://www.simplypsychology.org/likert-scale.html

McPartlan, P., Rutherford, T., Rodriguez, F., Shaffer, J. F., & Holton, A. (2021). Modality motivation: Selection effects and motivational differences in students who choose to take courses online. *The Internet and Higher Education*, *49*, 100793. https://doi.org/10.1016/j.iheduc.2021.100793

Merisotis, J. P., & Phipps, R. A. (1999). What's the Difference?: Outcomes of Distance vs. Traditional Classroom-Based Learning. *Change: The Magazine of Higher Learning*, *31*(3), 12–17. https://doi.org/10.1080/00091389909602685

Miller, J., & Ulrich, R. (2019). The quest for an optimal alpha. *PLoS ONE*, *14*(1), e0208631. https://doi.org/10.1371/journal.pone.0208631

Morales, C. R. (2017, August 24). *New CBRN requirement no issue for 773d CES*. Joint Base Elmendorf-Richardson. https://www.jber.jb.mil/News/News-Articles/Article/1287796/new-cbrn-requirement-no-issue-for-773d-ces/https%3A%2F%2Fwww.jber.jb.mil%2FNews%2FNews-Articles%2FNewsDisplay%2FArticle%2F1287796%2Fnew-cbrn-requirement-no-issue-for-773d-ces%2F

Nissen, M., McGuiness, R., & Davis, A. (2014). *Marine Forces Reserve: Accelerating Knowledge Flow Through Asynchronous Learning Technologies* (p. 55) [Technical Report]. Naval Postgraduate School.

Pappano, L. (2012, November 2). The Year of the MOOC. *The New York Times*. https://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html

Park, Y. J., & Bonk, C. J. (2007). Synchronous Learning Experiences: Distance and Residential Learners' Perspectives in a Blended Graduate Course. *Journal of Interactive Online Learning*, *6*(3). https://anitacrawley.net/Resources/Articles/Synchronous%20online%20course.pdf

Pienta, N. J. (2013). Online Courses in Chemistry: Salvation or Downfall? *Journal of Chemical Education*, *90*(3), 271–272. https://doi.org/10.1021/ed400097s

Reid, R. C., & Lambert, S. L. (2014). Using Contingency Planning Model to Mitigate the Impact of the Inevitable Disaster at the Class Level. *Information Security Education Journal*, *1*(2), 70–76.

Rosenberg, M. J. (2005). *Beyond E-Learning: Approaches and Technologies to Enhance Organizational Knowledge, Learning, and Performance*. John Wiley & Sons.

Seels, B. B., & Richey, R. C. (2012). *Instructional Technology: The Definition and Domains of the Field*. Information Age Publishing. https://books.google.com/books?id=fwIoDwAAQBAJ

Sharma, P. (2010). Blended learning. *ELT Journal*, *64*(4), 456–458. https://doi.org/10.1093/elt/ccq043

Smalley, A. (2021, March 22). *Higher Education Responses to Coronavirus (COVID-19)*. National Conference of State Legislatures. https://www.ncsl.org/research/education/higher-education-responses-to-coronavirus-covid-19.aspx

Stahle, L., & Wold, S. (1989). Analysis of Variance (ANOVA). *Chemometrics and Intelligent Laboratory Systems*, *6*, 259–272.

Stephenson, W. R. (2001). Statistics at a Distance. *Journal of Statistics Education*, *9*(3). http://jse.amstat.org/v9n3/stephenson.html

Strayer, J. F. (2012). How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learning Environments Research*, *15*(2), 171–193. https://doi.org/10.1007/s10984-012-9108-4

Strijbos, J.-W. (2011). Assessment of (Computer-Supported) Collaborative Learning. *IEEE Transactions on Learning Technologies*, *4*(1), 59–73. https://doi.org/10.1109/TLT.2010.37

Sun, J. C.-Y. (2014). Influence of polling technologies on student engagement: An analysis of student motivation, academic performance, and brainwave data. *Computers & Education*, *72*, 80–89. https://doi.org/10.1016/j.compedu.2013.10.010

Svrluga, S. (2020, April 7). Colleges are ditching letter grades this spring, but not all students are on board with 'ungrading.' *Washington Post*. https://www.washingtonpost.com/education/2020/04/06/college-grading-coronavirus/

Taking the GRE General Test at Home. (n.d.). ETS. Retrieved December 4, 2022, from https://www.ets.org/gre/test-takers/general-test/register/at-home-testing.html

Tanyel, F., & Griffin, J. (2014). *A Ten-Year Comparison of Outcomes and Persistence Rates In Online Versus Face-to-Face Courses*. 22.

Umpierrez, A. (2021, March 2). *Higher Ed Institutions Facing Increased Turnover Due to Burnout During Pandemic*. PLANSPONSOR. https://www.plansponsor.com/higher-ed-institutions-facing-increased-turnover-due-burnout-pandemic/

Vanslambrouck, S., Zhu, C., Lombaerts, K., Philipsen, B., & Tondeur, J. (2018). Students' motivation and subjective task value of participating in online and blended learning environments. *The Internet and Higher Education*, *36*, 33–40. https://doi.org/10.1016/j.iheduc.2017.09.002

Viorica-Torii, C., & Carmen, A. (2013). The Impact of Educational Technology on the Learning Styles of Students. *Procedia - Social and Behavioral Sciences*, *83*, 851–855. https://doi.org/10.1016/j.sbspro.2013.06.160

Ward, B. (2004). The Best of Both Worlds: A Hybrid Statistics Course. *Journal of Statistics Education*, *12*(3), null. https://doi.org/10.1080/10691898.2004.11910629

What Are the Different Types Of Educational Technology. (2021, July 31). *Edmonger*. https://edmonger.com/2021/07/31/types-of-educational-technology/

Wilson, S. G. (2013). The Flipped Class: A Method to Address the Challenges of an Undergraduate Statistics Course. *Teaching of Psychology*, *40*(3), 193–199. https://doi.org/10.1177/0098628313487461

Winquist, J. R., & Carlson, K. A. (2014). Flipped Statistics Class Results: Better Performance than Lecture over One Year Later. *Journal of Statistics Education*, *22*(3).

Wiyono, B. B., Indreswari, H., & Putra, A. P. (2021). The Utilization of "Google Meet" and "Zoom Meetings" to Support the Lecturing Process during the Pandemic of COVID-19. *2021 International Conference on Computing, Electronics & Communications Engineering (iCCECE)*, 25–29. https://doi.org/10.1109/iCCECE52344.2021.9534847

Zapalska, A., & Brozik, D. (2007). Learning styles and online education. *Campus-Wide Information Systems*, *24*(1), 6–16. https://doi.org/10.1108/10650740710726455

REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 074-0188*
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to an penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**			
1. REPORT DATE *(DD-MM-YYYY)* 24-03-2023	**2. REPORT TYPE** Master's Thesis		**3. DATES COVERED** *(From – To)* September 2021 – March 2023
TITLE AND SUBTITLE Student Performance in Traditional In-Person vs. Online Sections of an Introductory Mathematics Course			**5a. CONTRACT NUMBER**
			5b. GRANT NUMBER
			5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Kittle, Lauran E., Captain, USAF			**5d. PROJECT NUMBER**
			5e. TASK NUMBER
			5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology, Graduate School of Engineering and Management (AFIT/EN), 2950 Hobson Way, Building 640, WPAFB OH 45433-7765			**8. PERFORMING ORGANIZATION REPORT NUMBER** AFIT-ENC-MS-23-M-003
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Intentionally left blank			**10. SPONSOR/MONITOR'S ACRONYM(S)**
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **DISTRUBTION STATEMENT A.** APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.			
13. SUPPLEMENTARY NOTES This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.			

14. ABSTRACT

The growth of technology impacts nearly every aspect of everyday life, to include education and learning. However, it is imperative to consider exactly how the growth of technology impacts education. A detailed look into how student performance is affected by different classroom environments at AFIT-EN is conducted. This type of research, analysis, and modeling has yet to be conducted on a graduate school level. This study is completed on two data sets: the first is observational in nature with multiple demographic variables included which are then used to model student performance as determined by the student's final grade in the course. The second data set is experimental as the class was meticulously designed to eliminate as many sources of bias as possible, with the goal of making the in-residence and distance learning sections of the course mirror each other. Results indicated that when the sources of bias are controlled, either through variables or design of experiment, learning environment does not impact student performance. These results can be applied to other courses and schools, while also assisting professors in designing their courses and students in choosing which course section is right for them.

15. SUBJECT TERMS

Instructional Technology, Online, Distance Learning, In-Residence, Student Performance, Learning Environment

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Major Victoria Sieck, AFIT/ENC
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)*
U	U	U	UU	120	(509) 768-4853 (NOT DSN) (Victoria.sieck@afit.edu)

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18