2022

Success of Online and Face-to-Face Secondary Algebra I Students

Andrea Nicole Rohde
Walden University

Walden University

College of Education and Human Sciences

This is to certify that the doctoral dissertation by

Andrea Nicole Rohde

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Deborah Bauder, Committee Chairperson, Education Faculty
Dr. Wellesley Foshay, Committee Member, Education Faculty
Dr. Richard Hammett, University Reviewer, Education Faculty

Chief Academic Officer and Provost
Sue Subocz, Ph.D.

Walden University
2022

Abstract

Success of Online and Face-to-Face Secondary Algebra I Students

by

Andrea Nicole Rohde

MA, Christopher Newport University, 2007

BS, Christopher Newport University, 2006

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Education

Walden University

November 2022

Abstract

Research results from the past decade on the efficacy of online instruction for high school students varies. While many researchers suggested that younger students, mathematics students, and already struggling students who take online classes were underperforming compared to their classmates in the same course taken face-to-face (f2f), other researchers noted the advantages of online courses in providing flexibility, especially in pacing. Online education remains a popular course option for these learners despite the conflicting evidence of being able to truly support them in getting closer to graduation. Framed by Moore's transactional distance theory, a nonequivalent group quasi-experimental design was used to consider differences in student success and course completion between online and f2f environments for ninth-grade first-time Algebra I students while controlling for gender and ethnicity. Census sampling and logistic regressions were used to analyze archived data for 26,747 ninth-grade students who took Algebra I between 2016-17 to 2019-20 in the cooperating district's f2f or virtual courses, or the state-run virtual program. The independent variable consisted of the three environments while three dependent variables included (a) end-of-course final grades, (b) FastBridge aMath assessment results, and (c) completion as measured by course grade codes. Findings indicated that f2f students may have been outperforming online students, however, not all assumptions for analysis were met. This study may provide key information for improving both Algebra I and online education, increasing graduation rates, and better allocating funding in the educational budget.

Success of Online and Face-to-Face Secondary Algebra I Students

by

Andrea Nicole Rohde

MA, Christopher Newport University, 2007

BS, Christopher Newport University, 2006

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Education

Walden University

November 2022

Dedication

This dissertation is dedicated to the five most important people in my life, my husband and our four amazing girls. Brandton, your patience, sacrifices, and incredible dedication to this family is what has kept me afloat through the good, the bad, and the ugly…you will always be my reason not to touch the wall. To Cadence, Teagan, Kenna, and Brynn, I hope that someday you will understand that no matter how long things take or how hard it gets, if it’s worth doing, then it’s worth doing to the end. I love you all so much.

Acknowledgements

This endeavor would not have been possible without the help and support from Dr. Deborah Bauder. Words cannot express how incredibly grateful I am for your endless encouragement, patience, and advice throughout this writing process. I am also very appreciative for Dr. Wellesley Foshay who stepped up at the midway point to provide such valuable feedback and expertise on both the methodology and dissertation topic. I have learned so much from both you and I cannot thank you enough.

I would also like to thank Dr. Richard Hammett for your time and energy spent in reviewing my dissertation suggesting edits. You went above and beyond to include suggestions that made sure my paper was written to the highest standards at all times.

Finally, I would like to extend my sincere thanks to my past professors and classmates with whom I have shared this journey, and to the countless friends and family who have supported me in any and every way possible.

Table of Contents

List of Tables v

List of Figures vii

Chapter 1: Introduction to the Study..1

Background 6

Problem Statement 11

Purpose of the Study 14

Research Question(s) and Hypotheses...14

Theoretical Framework 16

Nature of the Study 17

Definition of Terms...18

Assumptions..22

Scope and Delimitations 24

Limitations 25

Significance...27

Summary 29

Chapter 2: Literature Review..30

Literature Search Strategy..31

Transactional Distance Theory 34

Transactional Distance Theory – Dialogue.. 36

Transactional Distance Theory – Structure.. 40

Transactional Distance Theory – Learner Autonomy.......................... 43

Limitations and Assumptions of Transactional Distance Theory 45
Transactional Distance Theory in Relatable Studies 46
Rational for Transactional Distance Theory Over Other Frameworks 50
High School Algebra I 55
High School Algebra I and Dialogue 55
High School Algebra I and Structure 61
High School Algebra I and Learner Autonomy 66
Key Variables, Designs, and Concepts 68
Comparing Online and F2f Algebra I Courses– Studies with Similar
Designs, Variables, or Research Questions 70
FastBridge Assessments as a Tool for Understanding Achievement 79
Comparing Online and F2f Algebra I Courses – Studies with Varying
Designs and Approaches 84
Summary 96
Chapter 3: Research Method...103
Research Design and Rationale 103
Methodology 107
Population 107
Sampling and Sampling Procedures 108
Procedures for Recruitment, Participation, and Data Collection 111
Operationalization of Constructs 111
Algebra I Course Environments... 112

Student Course Success 113

Student Standardized Assessment Success 113

Course Completion 114

Gender and Ethnicity 115

Data Analysis Plan 116

Threats to Validity 123

Ethical Procedures 130

Summary 133

Chapter 4: Results 135

Data Collection 137

Baseline Statistics 140

Analysis for RQ1 – F2f and District Virtual 144

Exploratory Analysis of RQ1 – F2F versus District Virtual 150

Analysis for RQ1 – District versus State Virtual 164

Analysis for RQ2 – F2F versus District Virtual 166

Exploratory Analysis RQ2 – Gender and Ethnicity Differences Within the F2F Environment 168

Analysis for RQ3 – F2F versus District Virtual 186

Exploratory Analysis of RQ3 – Gender and Ethnicity Differences for Course Completion 187

Summary 188

Chapter 5: Discussion 191

Interpretation of the Findings......192

Limitations......198

Recommendations for Future Studies......201

Implications for Social Change......204

Conclusion......205

References......207

Appendix......258

List of Tables

Table 1. 2016-2017 to 2018-2019 Consolidated Student Performance Comparison Summary ... 8

Table 2. Binomial Regression for F2F Versus District Dependent Variable Cumulative Splits ... 147

Table 3. Ordinal Logistic Regression Parameter Estimates F2F Versus District Virtual Course Grades ... 151

Table 4. Ordinal Logistic Regression Parameter Estimates Table for Non-Black/African American Versus Black/African American ... 153

Table 5. Ordinal Logistic Regression Parameter Estimates Table for Non-White Versus White ... 154

Table 6. Ordinal Logistic Regression Parameter Estimates Table for Non-Hispanic Versus Hispanic ... 155

Table 7. Ordinal Logistic Regression Parameter Estimates Table for Non-Asian/Multiracial Versus Asian/Multiracial ... 156

Table 8. Percentage of Students with EOC Letter Grades by Ethnicity and Gender Grouping ... 163

Table 9. District Virtual FastBridge ROI and Spring Raw Scores by Ethnicity and Gender ... 167

Table 10. Multinominal Logistic Regression Parameter Estimates F2F ROI Scores by Ethnicity and Gender ... 170

Table 11. Binomial Regression for Cumulative Splits on the DV for F2F Spring Raw Scores 175

Table 12. Multinomial Logistic Regressions for F2F Spring Raw Scores 179

List of Figures

Figure 1. Population Distributions for Different Learning Environments for ninth-grade Algebra I 142

Figure 2. Grade Distribution for Ninth-Grade Algebra I F2F Courses During the School Years 2016-17 to 2019-20. 158

Figure 3. Grade Distribution for Ninth-Grade Algebra I District Virtual Courses During the School Years 2016-17 to 2019-20. 159

Figure 4. Letter Grade Distribution for Ethnicity Groups in F2F and District Virtual Courses............ 160

Figure 5. Letter Grade Distribution for Gender Groups in F2F and District Virtual Courses............ 161

Figure 6. Number of F2F FastBridge Distribution Per ROI Group for Ethnicity........... 172

Figure 7. Number of F2F FastBridge Distribution per ROI Group for Genders............ 173

Figure 8. Percent of F2F FastBridge Spring Theta Score Distribution Per Ethnicity..... 177

Figure 9. Percent of F2F FastBridge Spring Theta Score Distribution Per Gender 178

Figure 10. Percentage of Students Who Did Not Complete the Course per Ethnicity and Gender............ 187

Chapter 1: Introduction to the Study

According to the Bureau of Labor Statistics (2017a, 2017b), the demand for students everywhere to graduate from high school to improve their chances of gaining acceptance to colleges or universities or earning higher-paying career opportunities has been increasing. The Occupation Finder of the *Occupational Outlook Handbook* by the Bureau of Labor Statistics (2020) showed that the 121 fastest-growing occupations all require at least a high school diploma, most of which also required on-the-job training. The days of finding a job that requires minimal education are beginning to dwindle. Instead, those in search of work and higher wages will most likely be required to have a high school degree, some type of postsecondary degree, and be more technologically savvy, including having digital problem-solving skills (Mamedova & Pawlowski, 2018; Organization for Economic Cooperation and Development [OECD], 2021). Between the years 2017 and 2020, laborers aged 25 to 64 with a tertiary degree had an employment rate of 81.9% compared to the same age group with anything below a secondary degree having an employment rate of only 55.1% (OECD, 2022). For those who do find themselves in the workforce, those with a lower level of education are more likely to be affected by decreased work hours and unemployment in times of national or global economic downturns such as was seen with the onset of the COVID-19 pandemic. OECD (2022) noted that the COVID-19 pandemic greatly affected the labor market and that individuals with lower levels of education had the number of hours worked reduced at a rate three times greater than those with higher levels of education. “Across the OECD, average hours worked fell by 8.5% among the highly skilled, 20% among those with a medium level of education, and 24% among those holding just a lower secondary education diploma or less: (OECD, 2021, section 1.3.2). Additionally, the OECD Skills for Jobs Database suggested that over the last 2 decades, “the demand for high-level cognitive skills has increased, while the demand for physical abilities and routine skills has decreased” (OECD, 2018, p. 20). Based on the statistics noted above, it would be reasonable to say that those who drop out of high school will most likely not be able to acquire the high level-cognitive skills necessary for the growing job market. Instead these dropouts may only have the skills necessary for labor intensive jobs which are trending downward, further limiting their career prospects.

Not only does dropping out of high school affect the student’s means of living and quality of life, but it also affects the rest of the nation. According to the Bureau of Labor Statistics (2017a), the rate at which high school graduates who did not enroll in college were either working or looking for work was 22% higher than their nongraduating cohorts. Additionally, of all the students who were working, high school graduates made, on average, $9,776 more per year than the dropouts (Bureau of Labor Statistics, 2017b). Rather than being able to provide for themselves, most high school dropouts, at some point, will have to rely on government assistance to meet basic living needs costing the American taxpayers approximately $292,000 over their lifetime (Graduation Alliance, 2017). Dorn et al. (2020) predicted that with the learning that was lost and high rates of dropouts due to the Coronavirus pandemic of 2020, the average high school student will lose $61,000 to $82,000 in lifetime earnings totaling an estimated $110 billion annually for that year’s graduating cohort. Additionally, according to Graduation Alliance (2017), students who dropped out of high school had a 4% higher unemployment rate than the national average, and high school dropouts accounted for over 80% of the nation's incarcerated population.

In the early 2000s, the U.S. government began a push to increase high school on-time graduation rates (Atwell et al., 2021). In 2001, the average freshman graduation rate was 71%, however, the U.S. government stepped in and created legislation that put school administrators and policymakers under increasing pressure to raise graduation rates to over 90% for each state by 2020 (Atwell et al., 2021). By the end of the school year 2019, only eight states had reached the goal of 90% high school graduation rate putting increasing pressure on the remaining 42 states to also reach that goal (Atwell et al., 2021). According to a report by the National Center for Educational Statistics (NCES, 2021a), during the 2019 school year, there were about 2 million 16 to 24-year-olds who had not completed high school and were not enrolled in school, an improvement from the 3.8 million in 2001 (NCES, 2021b). As a means of meeting the demand for getting students to pass required courses and graduate from high school, many school divisions across the nation began offering high school courses online, including Algebra I courses for first-time high school students. Two such programs offering online education courses are a state-run online program and a local school district in the current study's region. Since its start in 2013, the district's online program has been offering services to students in Grades 6-12 for both accelerated and credit recovery options (District Virtual, 2021). The state-level program has been in operation since 2005 and currently offers "125 unique core curriculum, AP, and elective courses with 281 variations, including SAT

Preparations" (Georgia Virtual Learning, 2014, para. 2), including credit recovery beginning in 2007.

In most high schools across the United States, Algebra I remains one of the core courses required to graduate (Education Commission of the States [ECS], 2019). In addition to course completion, most states require that end-of-course or exit exams also be passed (Achieve, 2016). During the 2018-2019 school year, one district of the study region reported that 26.1% of their high school Algebra I students failed their end-of-course state assessments and received an achievement level indicator of beginning learner (Governor's Office of Student Achievement [GOSA], 2020). According to the state department of education (DOE), "beginning learners do not yet demonstrate proficiency in the knowledge and skills necessary at this grade level…students need additional academic support to ensure success in the next grade level or course and to be on track for college and career readiness" (Georgia Department of Education [GADOE], 2022a, para. 3).

In March of 2020, the global pandemic that had been sweeping across the world, known as the Covid-19 virus or Coronavirus, immobilized the world's economy and social norms, including established educational norms (Kumar et al., 2021). In the wake of the epidemic, school systems everywhere were shut down to continuing face to face (f2f) classes as government requirements for social distancing were being mandated. To continue educational procedures, schools and school districts were forced to turn to online learning for all students. The United Nations International Children's Emergency Fund (UNICEF, 2020) estimated that nearly 90% of countries around the world turned to remote learning, but that 31% of students, nearly 463 million, could not be reached due to a lack of necessary technologies. By January 2021, most of the school districts in the United States moved to some form of hybrid learning plan which allowed students to be f2f with teachers at least 1 day a week or had plans to phase in hybrid and fully f2f return of students (Lieberman, 2020). As schools continued to try to make their way back to previous learning protocols with most students being back in school and f2f instruction, districts were facing harsh criticisms on both sides from parents, local health departments, labor unions, and teachers as to whether it was safe to return to school or better to maintain a remote/hybrid model (Gewertz, 2020). In some form or fashion, it seemed remote learning was here to stay, either as a completely remote model or a hybrid one. There are questions on the effectiveness of remote compared to f2f learning for all students, but particularly first-year high school Algebra I students. Research is needed to determine if one form of remote learning is more effective than the others and whether just one design should be used.

This chapter begins with a look at background research related to high school dropouts, online learning, and the risks of failing to make progress towards graduation for students who fail Algebra I in their first year of high school. A brief discussion of the problem statement, purpose, and research questions for this study is followed by an overview of Moore’s (1972) transactional distance theory, the theoretical framework that was used to guide this study. The final sections of this chapter include abbreviated rationales for the nature of this comparative study, definitions, assumptions, limitations, and the significance of this study.

Background

Algebra I has often been called the gateway to high school mathematics courses because a student's performance in Algebra I is often an indicator of how well they will perform in subsequent math courses (American Institutes for Research [AIR], 2017). Failing Algebra I during a student's first year of high school puts that student at risk of not graduating on time (AIR, 2017). According to Research Brief 6 of the American Institutes for Research (2017), "less than half of the students who failed Algebra I in ninth grade recovered the course credit by their fourth year of high school" (p. 3), and, of the students who were able to recover their Algebra I credit, only 62% were able to graduate on time.

According to the Office of Planning, Evaluation, and Policy Development (OPEPD, 2017), in addition to credit recovery programs, there are several programs in place to assist students in progressing towards graduation, some of which include early warning systems, academic tutoring, and case management. Each of these services and aspects have been, or are currently being, studied and scrutinized for the opportunities they provide for students to succeed (Alvarez & Marsal, 2018; Anderson, 2016; Malkus, 2018; Tromski-Klingshirn & Miura, 2017). Some of the current focus areas include "instructional methods, instructor certification…[and] program effectiveness" (Noble et al., 2017, p. 1), software for online learning (Lara et al., 2017), the use of blended learning as opposed to pure online learning (Barbour, 2017; Molnar et al., 2019), and the credibility of credit recovery courses (Barbour, 2017; "Grad Inflation; Education", 2019). Unfortunately, most of these studies seemed to focus on what to do after a student has failed rather than considering if the options of taking the course initially online or f2f had an impact on whether they succeeded in the first place.

Algebra I is arguably one of the most valuable mathematics courses covered in school as it is involved in nearly every aspect of everyday life, from being the essential building block in understanding other subjects such as social sciences and technologies; to home finances such as buying paint or buying a cell phone; to personal pastimes such as crafting and cooking (Cedar Tutoring Academy, 2018). As important as Algebra I is, it seems to be one of the more disliked courses for high school students and one of the most failed. Based on statistics from the GOSA (2020), over the past 3 school years (2016-2017 to 2018-2019), for the entire region, on average, 33.7% of students failed their Algebra I end-of-course state assessments. The next most failed end-of-course (EOC) assessments were Physical Science with a 38.5% fail rate, Coordinate Algebra with a 34.3% fail rate, and Analytic Geometry with a 33% fail rate (see Table 1 for all EOC assessment fail rates).

Table 1

2016-2017 to 2018-2019 Consolidated Student Performance Comparison Summary

End of course test	Annual fail rate (%) 2016-2017	2017-2018	2018-2019	*M*
Ninth-grade Literature and Composition	17.30	18.90	14.50	16.90
Algebra I	34.30	33.90	33.00	33.73
American Literature and Composition	19.70	21.20	20.40	20.43
Analytic Geometry	32.90	34.90	31.30	33.03
Biology	30.70	28.90	29.00	29.53
Coordinate Algebra	34.50	34.20	34.30	34.33
Economics/ Business/ Free Enterprise	27.80	24.10	24.00	25.30
Geometry	23.70	27.60	26.50	25.93
Physical Science	42.80	37.70	35.00	38.50
U.S. History	25.10	24.40	22.00	23.83

Note. Fail rates indicated the percentage of students identified as “Beginning Learners”. Average Pass Rates and Average Fail Rates were calculated using the data included in the table (GOSA, 2020). Adapted from “Downloadable data explained: Graduation and dropout rate” by Governor’s Office of Student Achievement, 2020, (https://gosa.georgia.gov/sites/gosa.georgia.gov/files/OBIEE-Help/Graduation___Dropout_Rate.htm).

Over the past several years, studies have shown that students, especially those who were already struggling, who took courses online at the postsecondary level did not fare as well as those who took the same course in a f2f classroom (Arias et al., 2018; Heppen et al., 2017; Protopsaltis & Baum, 2019). These studies suggested that students taking online courses were less likely to complete the course than students taking the same course in a f2f format (e.g., Arias et al., 2018; Oviatt et al., 2016; Protopsaltis &

Baum, 2019) and that students in online courses tended to perform more poorly than students in the f2f counterparts (e.g., Hart et al., 2019; Miron & Gulosino, 2016; Protopsaltis & Baum, 2019). Additionally, Belland et al. (2019) suggested that rushing through assignments or taking shortcuts, heavily relying on teachers or groups for answers rather than interacting with technology, and irregular attendance was more common in lower-level courses and credit recovery courses.

While most of these studies provide a look at what may have been happening at the postsecondary levels, they failed to give a clear picture of what was happening in high schools, especially for Algebra I students. Of the reports that considered the effectiveness of online courses, few were focused mainly on Algebra I students, and even fewer still compared online classes to f2f ones (e.g., Center on Innovations in Learning, 2015; Picciano et al., 2015). Additionally, several researchers have noted that most research surrounding online and f2f comparisons is riddled with validity issues, some more than others, and often should be interpreted with caution (Fryer & Bovee, 2018; Niu, 2020; Valverde-Berrocoso et al., 2020). Further suggestions for interpreting online research studies caution that often there are elements of online learning that are out of the researchers' control but can affect learning outcomes, such as teaching strategies, course design, student motivation, etc. (e.g., Cole et al., 2021; Martin & Bolliger, 2018; Valverde-Berrocoso et al., 2020).

Studies with similar concentrations as my study include those by AIR which was focused on the differences between online and f2f credit recovery Algebra I summer school courses (Heppen et al., 2017), Heissel (2016) who studied eighth-graders taking

Algebra I online or f2f; and Hart et al. (2019) who evaluated course performance and the likelihood of taking and passing follow-on courses for virtual and f2f courses for both initial attempts and credit recovery across several course subjects. While Heppen et al. (2017) and Heissel (2016) found that students in f2f courses tended to outperform students in online classes, Hart et al. (2019) found that students in online courses tended to perform better.

While research in the field is steadily growing, the contradictory findings of Heppen et al. (2017), Heissel (2016), and Hart et al. (2019) further demonstrated that there is still much debate among experts over the effectiveness of online learning, primarily noted as being due to the lack of quality research (e.g., Conway et al., 2016; Darling-Aduana, 2019; Hart et al., 2019; Mollenkopf et al., 2017; Molnar et al., 2019; Protopsaltis & Baum, 2019; Serdyukov, 2015). "Decades of research on the quality of online learning has yielded mixed results, which has led to confusion among educational administrators, educators, and students as to the efficacy of education in this format" (Mollenkopf et al., 2017, p. 2). Furthermore, there is a lack of quality quantitative data, especially for Algebra I students in online courses (Protopsaltis & Baum, 2019; Viano, 2018).

While there are several studies regarding Algebra I programs for f2f and distance learning, most have been focused on online learning platforms (e.g., Anderson, 2016), graduation rates for credit recovery students (e.g., Heppen et al., 2017; Powell et al., 2015), online credit recovery students' need for caring and community engagement (e.g., Barnett, 2016; Oviatt et al., 2016), games or apps for practicing mathematics (e.g., Farah et al., 2021; Marange & Adendorff, 2021; Umbara et al., 2021), or instructor and students' perceptions of the online environment (e.g., Everatt et al., 2019; Jaster, 2017; Niemi & Kousa, 2020; Oliver & Kellogg, 2015; Toker & Bektaş, 2021). When paired with Hart et al. (2019), Molnar et al. (2019), Protopsaltis and Baum (2019), and Serdyukov's (2015) statements regarding the lack of quality research for online learning, there stands a noticeable gap in the literature when it comes to an understanding of the relationships between the successes, both completion and pass rates, of high school Algebra I students attempting to earn credits for graduation in an online environment versus f2f.

Problem Statement

The problem I addressed was that students were being put into online Algebra I courses with little to no evidence as to the efficacy of online versus f2f courses for moving students closer to meeting graduation requirements (see Barbour, 2017; Molnar et al., 2019). "Currently, the most significant catalyst for dropouts is failure in Algebra I" (McGee et al., 2018, p. 2). The risks associated with dropping out of high school include being unemployed, living in poverty, needing public assistance (Powell et al., 2015), and at risk of having "different mental, social, occupational, and marital problems in adulthood" (Gubbels et al., 2019, p. 1). Furthermore, students who dropped out of high school were eight times more likely to be incarcerated than their graduating counterparts (Powell et al., 2015). In a 1987 report, the General Accounting Office stated that the social costs associated with increased high school dropout rates include decreased rates of skilled laborers and productivity and increased rates of public assistance and crime. As the interest and necessity in online learning continue to grow, understanding whether Algebra I online courses were helping students learn the skills necessary to graduate is an essential stepping-stone to fully grasping the potential of virtual courses. Getting more students to graduate not only helps them get into college or get better-paying occupations but also it affects the welfare of the economy by having fewer individuals unemployed and in need of taxpayer-funded government assistance (Dorn et al., 2020; Freeman & Simonsen, 2015; Heckman et al., 2018; Powell et al., 2015).

A consideration of the differences between the success of online Algebra I students and f2f Algebra I students can help politicians, school board members, and administrators make more informed decisions about virtual school policies and how to allocate resources, such as finances, training, time, or remediation towards improving the learning of high school Algebra I students as to not put them at risk of not graduating, especially since it is such a vital course in moving forward towards a degree. With such high stakes to society riding on whether students become dropouts and the risk factors of dropping out being highly dependent on earning credits towards graduation, then it seems only reasonable to verify that online Algebra I courses are moving students in the right direction towards graduation and equipping students with the skills necessary to be successful in the workforce (Brubacher & Silinda, 2019; Darling-Aduana, 2019; Dupere et al., 2015; Hart et al., 2019).

Over the past 2 decades, the number of studies concentrating on online learning has grown considerably (Heppen et al., 2017; Viano, 2018). However, most of these studies were centered on postsecondary education (e.g., Amparo et al., 2018; Arias et al., 2018; Cavanaugh et al., 2016; Chisadza et al., 2021; Conway et al., 2016;). Recent trends in high school educational research have been to evaluate the possible uses for online education, student and teacher perceptions and satisfaction in online learning, and communication techniques (e.g., Thai et al., 2020; Viano, 2018). Although Algebra I courses may be vital to a student's chance of graduating high school, research shows that most online learning is used for credit recovery, after a student has already failed a course (e.g., Clements, Stafford, et al., 2015; Clements, Zweig, et al., 2015; Darling-Aduana, 2019; Heinrich et al., 2019). Very little research exists on the effectiveness of online learning for students taking Algebra I for the first time (Barbour, 2017; Clements, Zweig, et al., 2015; Heppen et al., 2017; Hughes et al., 2015; Molnar et al., 2019; Powell et al., 2015; Viano, 2018). Additionally, of the studies focusing on high school virtual courses, there is some debate over online learning potential and actual learning outcomes (Hart et al., 2019). While several studies suggested that online courses have the potential to allow for flexible, individualized learning and can provide interactive resources to motivate students (Heppen et al., 2017; Kim et al., 2014; Viano, 2018; Wheatley, 2016), others suggested that students who are at-risk of failing or in need of credit recovery lack the reading and math skills necessary to take a technology-enhanced course (Adams, 2020; Hart et al., 2019; Molnar et al., 2019; Protopsaltis & Baum, 2019; Viano, 2018).

Even as online learning continues to grow at astounding rates, especially with the current pandemic, the true power it has for aiding Algebra I students is still unknown. What is still not known is the quality of the online Algebra I courses compared to f2f (Clements, Stafford, et al., 2015; Clements, Zweig, et al., 2015; Hart et al., 2019), student motivations for taking online courses (e.g., Hart et al., 2019), or how Algebra I online compares to f2f courses for terms longer than a 3-4 week summer term (e.g., Heppen et al., 2017; Hughes et al., 2015). My study is intended to help provide insight into these still unknown areas of online education.

Purpose of the Study

The purpose of this quantitative study was to assess and compare learning outcomes for one U.S. region's students who took an Algebra I course in a district-run virtual program, a state-run virtual program, or in a f2f environment to determine the efficacy of the two approaches for moving these students closer to meeting graduation requirements. A nonequivalent group quasi-experimental design was used to empirically evaluate whether there were differences in the success and completion between two different online groups and a f2f group.

Research abounds that considers the differences in student academic achievement based on geographic location, disadvantaged (socioeconomic) status, race, ethnicity, and gender (e.g., Chisadza et al., 2021; Cooper et al., 2019; Curtis & Werth, 2015; Protopsaltis & Baum, 2019). Therefore, student characteristics must also be considered when comparing first-time Algebra I students (Conway et al., 2016; Gates, 2020; Hart et al., 2019). Student demographics were considered as a covariate for the independent variable.

Research Question(s) and Hypotheses

RQ1: Is there a difference in student course success, as measured by end-of-course grades, between local online courses, state-run online courses, and f2f instructional environments in Algebra I courses, while controlling for student demographics, such as gender, ethnicity, and race?

$H_0$1: There is no difference in student course success between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

H_a1: There is a difference in student course success between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

RQ2: Is there a difference in student state assessment success, as measured by SOL assessment scores, between local online courses, state-run online courses, and f2f instructional environments in Algebra I courses, while controlling for student demographics, such as gender, ethnicity, and race?

$H_0$2: There is no difference in student course success between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

H_a2: There is a difference in student course success between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

RQ3: Is there a difference in course completion (as measured by course grade codes) between local online courses, state-run online courses, and f2f instructional environments, while controlling for student demographics, such as gender, ethnicity, and race?

$H_0$3: There is no difference in course completion between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

H_a3: There is a difference in course completion between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

In summary, the dependent variables of student success were measured by using students' final course grades, spring FastBridge aMath assessment scores, and rate-of-improvement (ROI) between assessments. The dependent variable of course completion was measured by students' end-of-course grade codes where a grade of A, B, C, or F indicated that the student has completed the course and a grade code of I (withdraw) indicated that the student did not complete the course. More information regarding the variables and how they were measured is discussed in Chapter 3.

Theoretical Framework

In this study, I used Moore's (1997) transactional distance theory (TDT) as the lens through which the online and f2f learning environments were compared. TDT is a means of describing the potential separation between learners and teachers that is not necessarily geographical. When there is a misunderstanding between teachers and learners due to the pedagogy or content, there is a transactional imbalance, making transactional distance a factor in both f2f and online education. No matter the instructional environment there are three components that contribute to the degree of transactional distance: the structure or design of the educational program, platform, or course; the amount and quality of teacher and learner dialogue; and the learner's amount of drive and intrinsic motivation to take control of their own learning, called learner autonomy (Moore, 1997). Moore suggested that a physical environment has a transactional distance between teacher and learner due mainly to cognitive differences, though communicative challenges also exist. In distance education (originally referring to correspondence courses), cognitive and communicative schisms are much greater, leading to greater transactional distances (Moore, 1997). Should f2f students outperform online students and have higher completion rates, TDT may be used to explain possible differences due to the likely lesser amount of transactional distance and would be consistent with prior research. Should online students perform better, further research would need to be conducted to investigate the impact of transactional distance on virtual learning outcomes for this district. Using the concepts presented in TDT, it is possible to compare learning environments and establish a more grounded study.

Nature of the Study

This quantitative research study used census sampling through archived data, and a nonequivalent group quasi-experimental design as it was intended to be used as a means of discovering the extent of the differences that instructional environments (IV) had on student course success (DV), standardized assessment success (DV) and course completion (DV) for high school Algebra I. Quasi-experimental research methods are most appropriate when the creation of treatment and control groups cannot be done through randomization (Ersen et al., 2018; White & Sabarwal, 2014), and group comparison approaches allow researchers to compare the dependent variable outcome results for distinct groups. A nonequivalent group quasi-experimental design was the most fitting approach for this research as the main goal was to compare educational outcomes for three groups of students that could not be created through randomization. While Milestones assessments scores were originally proposed as the standardized assessments, FastBridge aMath scores and their accompanying ROI were received and used to understand standardized assessment success instead.

As mentioned above, student course success, standardized assessment success, and course completion served as dependent variables, while the independent variables were the three learning environments of either local online, state online, or f2f. Additionally, student demographics, including gender, ethnicity, and race, served as covariates.

Both the online and f2f comparison groups were formed before the start of my study. For this reason, I obtained all data from the school board office archives, and I had no direct contact with any student or teacher. Once collected, regression analysis was conducted using Statistical Package for the Social Sciences (SPSS) software when possible. If not possible, data descriptions were provided. More information regarding quasi-experimental designs, the variables, and time/resource constraint considerations is discussed in Chapter 3.

Definition of Terms

At-Risk Students: At-risk students have a higher-than-average probability of dropping out or failing high school due to living circumstances (such as living in poverty, group homes, high minority areas, or being incarcerated); those who are in danger of academic failure due to insufficient credits for their grade level (Powell et al., 2015; U.S. Department of Education, 2007).

Course Completion: Students who remain in a course until the end of the term and receive a final course grade (A, B, C, D, or F) have completed that course (Los Angeles Southwest College, 2019; The Research and Planning Group for California Community Colleges, 2011).

Course Dropout: A course dropout is differentiated from Course Withdrawal as a student who stops attending a course or withdraws from a course after the withdrawal period has ended, prior to receiving a grade (GADOE, 2011; Sarikas, 2015).

Course Success: Course success means that a student has remained in a class until the end of the term, has received a passing grade of A, B, C, or D, and has met all further requirements as set forth by the course description (The Research and Planning Group for California Community Colleges, 2011). A student who receives a grade of F has completed the course but was unsuccessful at passing it.

Course Success Rate: The course success rate is a percentage calculated by dividing the number of students who successfully passed a course by the course enrollment number at the beginning of the semester and then multiplying by 100. Students who formally withdrew or dropped out after the add/drop period or did not complete the course are included in this calculation (The Research and Planning Group for California Community Colleges, 2011).

Course Withdrawal: A student has withdrawn from a course if he or she leaves the course before the designated add/drop period ends. The add/drop period ends after the 10th day from the first day of the course (GADOE, 2011).

Credit Recovery: Credit recovery refers to the attempt to retake a previously failed course and passing it allows them to earn back the necessary credits needed to graduate from high school (Georgia Credit Recovery Program, 2016; Malkus, 2018).

Ethnicity/Race: GADOE (2020a) differentiates between the ethnicity and race of a student. The ethnicity of a student is defined as being either Hispanic or Non-Hispanic, but not both, whereas a student's race is defined as Indian, Non-Indian, Asian, Non-Asian, Black, Non-Black, Pacific, Non-Pacific, White or Non-White. The data received from the district did not have the same distinction as the state DOE. In the provided data set, the title heading was federal ethnicity code, and the available options were White, Black/African American, Hispanic, Asian, Multiracial, American Indian or Native Alaskan, or Native Hawaiian or Pacific Islander. For the remainder of this study, ethnicity and race will be labeled as just ethnicity and will be in accordance with the district-provided federal ethnicity codes.

F2f Classroom: A f2f classroom is one in which instruction, assignments, and assessments are delivered in person through f2f contact between students and instructors. (Purdue University, 2017).

FastBridge Data: FastBridge analysis measures the amount of learning growth a student has achieved from one quarter to the next. This ROI is calculated by finding the difference between 2 quarters' assessment scores and dividing it by the number of weeks between the two assessments (Brown, 2021). Since the cooperating district supplied spring raw and ROI scores instead of the Milestones scores that were expected, raw scores and ROI were used to investigate RQ2. More on FastBridge assessments is in Chapter 4.

Georgia Milestones: The Georgia Milestones were the end-of-course state assessments for specified high school courses that were meant to measure “how well students are mastering the state-adopted content standards” (GADOE, 2020b, para. 2), or the content-specific Georgia Standards of Excellence. Beginning in the 2014-2015 school year, the Georgia Milestones replaced the End of Course Tests (GADOE, 2020c). Additionally, the scores achieved on the Georgia Milestones count for 20% of a student’s end-of-course grade (GADOE, 2020b). Georgia Milestones were expected to answer RQ2, however, FastBridge aMath assessment scores were received instead. More on this change is discussed in Chapter 4.

Georgia Standards of Excellence (GSE): The Georgia Standards of Excellence (GSE) were the state assessment standards that began replacing the Georgia Performance Standards (GPS) as updated standards during the 2015-16 school year. The GPS were a means to “provide clear expectations for instruction, assessment, and student work. They define the level of work that demonstrates achievement of the standards” (GADOE, 2020d, para. 2.).

Graduate: A student is considered a graduate if he or she has earned a diploma that has been approved by the Board of Education (GADOE, 2020a). “Students who drop out of school and receive a GED certificate are not considered high school graduates" (GADOE, 2012, p. 2).

High School Dropout: A high school dropout is a student who stops attending school prior to earning a high school diploma (GADOE, 2012).

Online/ Virtual Classroom: "Online education is defined as education being delivered in an online environment through the use of the internet for teaching and learning…[including] online learning on the part of the students that is not dependent on their physical or virtual co-location. The teaching content is delivered online, and the instructors develop teaching modules that enhance learning and interactivity in the synchronous or asynchronous environment" (Singh & Thurman, 2019, p. 302). This allows teachers and students to be in geographically separate locations. The state DOE specifies that online courses can be indicated either by course number by placing a 3 immediately following the decimal (xx.3) or by with the ONLINE COURSE indicator set to 'Y' (GADOE, 2020e).

Student Enrollment: Student enrollment is a total count of students enrolled in all courses, duplicating the count of students enrolled in more than one course (Los Angeles Mission College, Office of Institutional Research and Planning, 2011; University of Northern British Columbia, 2012).

Assumptions

Making comparisons between two groups holds the greatest value if the subjects have mostly commonalities and only very few differences (Child Care and Early Education Research Connections, 2019; What Works Clearinghouse [WWC], 2017a;

White & Sabarwal, 2014). For this study, I assumed that the content, rigor, and evaluation methods of the f2f Algebra I courses were equitable to the content, rigor, and evaluation methods covered in the virtual Algebra I courses. This included assuming that the students were administered the same test, or that tests were equated to use the same scale. These assumptions were necessary as it was not possible to evaluate every aspect of both environments while also trying to compare the student outcomes in each.

Since all three environments were using the same standardized assessment, it was adequate to assume the groups' testing outcomes were congruent. By assuming the three environments to be compatible in curriculum content, pedagogy, pacing, and assessments, the study's focus could be centered on the success and completion of the students in each environment instead of laying too much concern into whether one environment had it "easier" than the other or not. Also, assuming the content, rigor, and evaluation proceedings were similar allowed for Moore's TDT (1972, 1993) to become the baseline by which the environments could be compared. However, it must also be noted that rarely is it the case that online courses precisely match up with their f2f counterparts. This is discussed in further detail in the limitations section.

The original proposal for this study was to use a quasi-experimental comparison-group design with the assumption that online and f2f students would have adequate similarities at the beginning of the study to contribute differences in outcomes to the learning environments. Additionally, a lack of end-of-course grades or assessment scores from the previous year meant there was no way to empirically assess the differences in the online and f2f groups prior to the Algebra I courses, and that the assumption had to be made that they were equivalent as far as base knowledge. This type of assumption makes the study more vulnerable to threats to internal validity. Limitations of this type of design and assumptions are discussed later in this chapter.

Scope and Delimitations

The original focus topic and locale for the study were initially chosen out of my personal experience and insight into both the f2f and online environments while teaching high school geometry at a school within the originally proposed school district. However, after suffering a damaging fire to one of the local high schools and then being subject to the educational chaos following the shutdown of schools due to the Coronavirus epidemic, the originally proposed school district was unavailable for research. During the COVID-19 spread across the nation, many schools were forced to turn to online learning as the main outlet for education. While most school districts across the nation rejected research inquiries while they dealt with finding ways to ensure quality education for their students, one school district acknowledged the need for data-driven evidence to support their educational decision making. However, instead of looking at the success and outcomes of the effectiveness of online and f2f environments for geometry credit recovery students, this district requested results for the success of online and f2f students taking Algebra 1, including a comparison of the students in their own virtual program versus the state-run program.

After delving into the literature, I realized that there was little to no research that specifically compared the learning outcomes for online and f2f Algebra 1 students, hence the need for this study. I focused solely on high school students from the new study district taking Algebra 1 courses. My study did not include students from outside of the specified district, students who were in blended learning courses (partly f2f and partly online), or students who took Algebra 1 credit recovery courses.

Due to the state's specifics regarding required credits for graduation and the unique population of the study region, the results of this study are not generalizable to all high school online programs. However, administrators of such programs can use the results of this study to be better informed as to possible outcomes for their students.

Limitations

As noted earlier, using a nonequivalent quasi-experimental design and operating under the assumption that students had equivalent prior mathematical knowledge created severe limitations on how the results of this study could be interpreted. Using FastBridge ROI growth scores and data cleaning techniques helped to mitigate possible baseline differences between groups allowing for a discussion that the differences in outcomes were most likely due to the differences in the independent variable. However, results had to be considered with caution as FastBridge data was only available for one of the 4 years requested.

The two virtual high school programs that were considered for the study were a division-specific virtual program and the state-run online program. The division-specific virtual program was for district students only while the state-run virtual program offered courses to students from all over the state. To establish similarities between the two online groups and the division f2f group, only students from within the district were considered for this study. This may have limited the generalizability of the study but helped to maintain baseline comparisons.

For online virtual courses to be worthwhile, they must ensure the students are truly learning the material necessary to pass the EOC assessments and be successful outside of the classroom instead of just presenting them with a passing grade (Powell et al., 2015). Some educators and researchers claim there are several credit recovery and virtual programs that lower their standards for scope and rigor so students can pass the course (Malkus, 2018; Powell et al., 2015). This study was limited in that I did not present any judgments one way or another as to whether the online courses involved were equitable to the f2f courses in maintaining standards and providing genuine learning opportunities through meaningful interactions with peers, teachers, and content. Additionally, I did not attempt to make comparisons regarding the depth and breadth of the covered course content between the three environments. Instead, the study was completed under the assumption that the environments were, in fact, equivalent, as noted above. Evaluating the comparability of content and rigor of the online Algebra I courses for this region may be pursued in future studies but was beyond the scope of the present one. Without a deeper understanding of how the three environments compare as far as content and rigor, this study was strictly limited to be a foundation upon which further studies can be conducted.

True experimental research designs randomly place participants into control and treatment groups (McKinley & Rose, 2019). For this study, it was unclear how students were placed in courses, whether it be due to student preference, scheduling conflicts, health reasons, administrative placements, or other reasons. However, it was assumed that students were not placed strictly by chance. The inconsistent method by which students were placed into the three types of classrooms was considered a limitation. To best provide a reliable study with this limitation, census sampling was used. This is discussed further in Chapter 3.

Yet another limitation of the study was in the use of the selected dependent variables. The use of student course grades as a dependent variable was potentially hazardous as these grades were generally subjective scores that were not standardized and most likely included some measurement errors. However, including the use of a second standardized dependent variable, FastBridge aMath assessment scores, allowed for a stronger, more substantial study. Further discussion regarding the variables is in Chapter 3.

Finally, several studies have mentioned that student demographics, such as socioeconomic status, can play a role in predicting the success of student achievement (e.g., Cavanaugh & Jacquemin, 2015; Cooper et al., 2019; Gates, 2020; Hart et al., 2019; Protopsaltis & Baum, 2019). Unfortunately for this study, socioeconomic status was not an available data point, and therefore could not be used as a covariate. This limited the study in that it was impossible to specify whether student performance was affected by low or high family income.

Significance

While studies regarding online environments have been conducted that point to the possible advantages for student learning (e.g., Barnett, 2016; Curtis & Werth, 2015;

Ersen et al., 2018; Heppen et al., 2017; Malkus, 2018; Oliver & Kellogg, 2015; Serdyukov, 2015), there was a lack of evidence that demonstrated the effectiveness for first-time Algebra I students (Barbour, 2017; Clements, Zweig, et al., 2015; Heppen et al., 2017; Hughes et al., 2015; Powell et al., 2015; Viano, 2018). In studying the completion and success of high school Algebra I students in online and f2f environments, I attempted to provide new evidence as to whether online learning is as effective as f2f for these students. Results from this study could provide more insight for school administrators and counselors to place students in an environment best suited to individual needs. Additionally, using TDT as a framework for comparing online and f2f outcomes could provide further support upon which future research that considers factors of interactivity, communication, isolation, and motivation for ninth-grade math students may be based.

The potential for social change based on the results of this study may also provide key information for improving Algebra I education, increasing high school graduation rates, decreasing dropout rates, and better-allocating funding from an already tight educational budget. As stated earlier, getting more students to graduate not only provides the students with more opportunities to be successful but also provides relief for society by having fewer individuals unemployed and in need of tax-payer-funded government assistance (Dorn et al., 2020; Freeman & Simonsen, 2015; Goodwin et al., 2016; Gubbels et al., 2019; Heckman et al., 2018; Horton, 2015).

Summary

High school Algebra I, often called the gatekeeper course, remains one of the core courses required to graduate throughout most of the United States (AIR, 2017; Rickaby, 2021). However, many students are failing this course which puts them behind in completing follow-up courses for graduation, and are then required to complete credit recovery courses, either online or in a f2f environment, to attempt to make up the credit (Hart et al., 2019; Heppen et al., 2017). McGee et al. (2018) stated that Algebra I was the "most significant catalyst for dropouts" (p. 2), putting these students at risk of not graduating high school. Of the previous research conducted, few, if any, make worthy comparisons between online and f2f environments, especially between local and state-run virtual programs. The coming chapter provides an overview of past research regarding Algebra I courses and online learning, as well as an in-depth analysis of how Moore's (1993) TDT was used to evaluate the course completion and success of high school Algebra I courses for online and f2f students.

Chapter 2: Literature Review

The problem addressed was that students were being put into online Algebra I courses with little to no evidence as to the efficacy of online versus f2f courses for moving students closer to meeting graduation requirements (Barbour, 2017; Molnar et al., 2019). The purpose of this quantitative study was to assess and compare learning outcomes for one U.S. region's students who took an Algebra I course in a district-run virtual program, a state-run virtual program, or in a f2f environment to determine the efficacy of the two approaches for moving these students closer to meeting graduation requirements. The NCES (2015) reported that during the 2009-2010 school year, there were just over 1.8 million students in the United States enrolled in one or more distance learning courses, but by the 2019-2020 school year, that number had increased to over 50.4 million (NCES, 2020) and continues to expand. The rapid growth of online education programs has been mostly attributed to their alleged roles in aiding in the development of increased student learning achievements and saving money for the school districts that employ them (Hughes et al., 2015; Martin & Kumar, 2018; OPEPD, 2017). However, much of what research there is regarding online education tended to focus on multiple areas throughout the realm of education with little, if any, being concerned with Algebra I courses and online learning for first-year high school students, specifically inquiring as to the effectiveness of the district-level virtual, state-run virtual, and f2f programs. This chapter provides a summary of the theoretical foundations through which the study was viewed and provides a discussion of the past research related to Algebra I courses and online learning.

Literature Search Strategy

Most of the articles used for this study were found using the ERIC database, Education Source database, Georgia Department of Education website, and the U.S. Department of Education website. Initial searches were for terms that covered a broader spectrum of topics, such as *online learning*, *Algebra I students, dropouts, secondary students, learning theories,* and *mathematics*. Each search resulted in a massive variety of topics that had to be narrowed to fit the focus of this study. The final count of reviewed articles was just over 700. Approximately 300 of them were cited.

Searching for *online learning* resulted in a list of over 40,200 articles between ERIC and Education Source databases. To narrow search results without losing highly relevant articles, search filters were set to include articles with related words, such as *virtual learning* and *distance learning*, but only articles that had been peer-reviewed were published between 2015 and 2022 and were based on high school or secondary students as participants or immediate stakeholders. Of the 10,794 resulting articles, many were focused on student attitudes, perceptions, and motivation; teacher preparation classes or professional development courses for teachers at secondary schools; COVID-19; or the use of social media in the classroom. To further filter results, search parameters were refined to include *online learning AND high school OR secondary education AND online versus f2f AND mathematics*; peer-reviewed articles only; and published between 2016 and 2022. These specifics still resulted in 1,146 studies. After additional amending of search parameters, only articles related to perspectives and experiences of online learning, best practices for teaching online courses, reviews of effective and noneffective online learning programs, online and f2f comparisons, and articles concerning online mathematics courses were considered. A total of 161 articles were initially reviewed, while 104 articles were included for this aspect of the report.

Focusing on particular groups of students required more specific terms to be investigated. The terms *at-risk students, dropout*, and *Algebra I* were each examined individually and as a combined search term using both AND OR. *At-risk* alone turned up over 164,256 articles but included those related to medical and financial risks. Parameters for at-risk were then changed to *at-risk students*, resulting in 29,807 results. When using the search parameters of *at-risk students AND Algebra I* only 21 articles were returned. Only one article was both peer-reviewed and written since 2017, but its focus was on virtual manipulatives for students with learning disabilities. *Online learning AND Algebra I* returned four articles, only one of which was peer-reviewed and written in the last 5 years. This article, by Satsangi et al. (2021) was also focused on virtual manipulatives for Algebra I high school students with learning disabilities.

To ensure relevant articles were not being excluded, I conducted more general searches using just *at-risk students* and just *Algebra I*. These searches were filtered by including only the articles that had been peer-reviewed, were written between 2015 and 2022, and included varying combinations of linked subjects. Several combinations of linked subjects were used to attempt to find the articles with the most relevant subject matter without eliminating those that could still be useful to the study. One such combination included high school students and secondary as linked subjects to "Algebra I". This filter combination resulted in 143 articles, but only 62 were peer-reviewed. Only

22 of these articles were published since 2017 and peer-reviewed. After nearly 20 search combinations were performed and filtered, a total of 403 articles were reviewed with 212 proving at least minorly relevant to the current study.

When researching TDT, two major searches were conducted using the ERIC and Education Source databases. First, a general search for TDT using Education Source resulted in 175 articles that had been peer-reviewed. Initially, a filter was set to include only articles published since 2017 which resulted in 74 articles. However, several of these articles were related to higher education, therapy, or medical school. A second search was conducted with search terms *Algebra I OR secondary school mathematics* and transactional distance theory resulting in zero articles. The original search of just *transactional distance theory* was reset and the filter for more recent years was then removed for this search so that articles written by the original author were not excluded.

The same search in the Thoreau database resulted in 483 articles. Initially, these were filtered down to include only those written within 5 years of the time of this writing resulting in 207 articles. These were each scanned for relevance to the topic of the current study prior to being considered for extensive review. When searching *transactional distance theory AND secondary school*, without filtering for peer-reviewed articles or published within the last 5 years, only four articles were returned. A search for *transactional distance theory* and *mathematics* was conducted with no filters for dates or peer-reviewed articles. Only seven articles resulted, each of which was a duplicate from the ERIC and Education Source database searches. Eventually, articles prior to the last 5 years were added for review, and additional search terms were removed. A final listing of over 20 articles from Thoreau and 40 from ERIC/Education Source with mid to high relevance to the topic of this study was reviewed regarding TDT. After including articles relevant to methodology, including terms such as *quantitative research, sample size,* and *logistic regression*, over 300 articles were reviewed and cited for this study.

Transactional Distance Theory

Moore is considered one of the leading theorists of online and distance education (Garrison, 2000; Giossos et al., 2009; Grigoryan, 2017; Krieger, 2017; Larkin & Jamieson-Proctor, 2015; Murphy & Rodriguez-Manzanares, 2008; Weidlich & Bastiaens, 2018; Wengrowicz & Offir, 2013). Moore's TDT stems from Dewey's (1939) theory of transaction and was first termed in 1972 (Dockter, 2016; Giossos et al., 2009; Krieger, 2017; Moore, 1997; Murphy & Rodriguez-Manzanares, 2008) when many colleges and Army courses were in the beginning phases of what is now known as distance education (previously known as correspondence classes). As technology has evolved, so has Moore's theory. Yet, TDT remains one of the founding theories for distance education, particularly its ability to explain the impact of student learning in distance education courses (Aluko & Omidire, 2021; Duc, 2012; Elyakim et al., 2019; Garrison, 2000; Grigoryan, 2017; Krieger, 2017; Larkin & Jamieson-Proctor, 2015; MacLeod et al., 2019; Martin & Kumar, 2018; McMillion, & King, 2017; Murphy & Rodriguez-Manzanares, 2008; Swart & MacLeod, 2021; Tirronen et al., 2020; Weidlich & Bastiaens, 2018; Yilmaz, 2017; Yilmaz & Keser, 2017).

As suggested by several authors, in its infancy, TDT was applied to correspondence classes (the mailing of study materials back and forth between teachers and students) which dominated the field of distance education as personal computers were not readily available (Huang et al., 2015; Kassandrinou et al., 2014; Krieger, 2017; Moore, 1989, 1997). Moore (1989, 1993, 1997) specified that most of the educational materials of that time consisted of instructional booklets and videos recorded by the instructor and sent through post to the student. Once materials arrived at the student's location, it was up to the student to decipher the information and decide the message's intent. If students had questions regarding the material they would have had to mail their concerns to the instructor and patiently wait for a reply which could have taken weeks to come (Moore, 1989, 1993, 1997). The waiting may have led to a potential misunderstanding of the subject and material, misusing it, or not using it to its full potential (Dockter, 2016; Huang et al., 2015; Kassandrinou et al., 2014; Moore, 1989). What Moore (1993) originally proposed was that it was not the physical distance between the teacher and student that made learning difficult, but the lack of communication between the two.

The main purpose of Moore's theory is to help define and describe the psychological and communicative disconnect between teacher and student caused by the lack of interaction occurring with these types of courses (Huang et al., 2015; Kassandrinou et al., 2014; Moore, 1989, 1993; Weidlich & Bastiaens, 2018), and to distinguish the types of interaction that are necessary to provide successful distance learning environments (Moore, 1997). As Moore's theory began to take shape, it became apparent that the psychological and communicative separation between teacher and student was not just happening between those who were geographically separate, but also between teachers and students who were in the same room, though the time lag in distance education did cause greater transactional distance.

In the initial stages, Moore (1993, 1997) suggested there were two elements at play in determining transactional distance: structure and dialogue (Swart & MacLeod, 2021). Later, learner autonomy was added as a third, but equally vital, factor (Huang et al., 2015; Moore, 1993, 1997). Each of the three elements of Moore's, dialogue, structure, and learner autonomy, are considered qualitative variables that are ever-changing and flowing and are different for every learner and every situation. Moore (1997) suggested that the amount of transactional distance experienced by a learner varies with the changes in the quantity and quality of dialogue, the amount of flexibility of the learning structure, and the ability of the learner to be self-driven. In the following sections, each of the three elements is discussed in detail, including their impact on transactional distance.

Transactional Distance Theory – Dialogue

In the initial stages of Moore's (1997) TDT, dialogue was essentially defined as the communication that occurred between the instructor and the student. In the earlier correspondence courses, video recordings, self-study materials, and mail by post dominated the medium through which these communications took place, leading to much higher transactional distance. As technology advanced, the medium through which students could communicate with their instructors (audioconferencing, for example) became more interactive and more immediate, resulting in a lower perceived transactional distance (Moore, 1997). Eventually, students became able to communicate with their instructors and each other through audio conferencing and email (Moore, 1997). Hence, one of the earlier changes to the definition of dialogue was to include exchanges between instructors and exchanges with peers.

Moore (1997) suggested that the term dialogue is often used synonymously with the term interaction but noted that there is a significant difference between the two. Interactions can be one-way, negative, and detrimental to a learning environment while dialogue is an interaction or series of interactions that are reciprocal, positive, constructive, purposeful, and used to improve learning (Elyakim et al., 2019; Grigoryan, 2017; Larkin & Jamieson-Proctor, 2015; Moore, 1997; Quong et al., 2018). The updated definition meant that the overall dialogue that occurs between student and instructor or student and student should be fashioned in such a way that it is constructive and valuable to all parties involved.

Most researchers seem to agree on the theoretical definition of dialogue, however, as education has spread itself into the realms of internet use, social media, discussion boards, videoconferencing, and more, the development of an operational definition for dialogue continues to be debated (Ekwunife-Orakwue & Teng, 2014; Giossos et al., 2009; Quong et al., 2018; Wengrowicz & Offir, 2013; Zhang, 2003). Most researchers agree with Moore’s (1972, 1993, 1997) original definition that there are three distinct forms of dialogue, learner-content, learner-instructor, and learner-learner. Learner-content interactions are considered the most basic of interactions, those in which the student reads, watches, listens to, or learns the material without mediation from an instructor (Duc, 2012; Ekwunife-Orakwue & Teng, 2014; Elyakim et al., 2019; Hillman et al., 1994; Quong et al., 2018; Wengrowicz & Offir, 2013; Zimmerman, 2012). In distance learning courses with limited technology, this could mean that the student would need to have an internal dialogue about the course material making the dialogue highly one-sided but considered dialogue nonetheless (Duc, 2012; Elyakim et al., 2019; Moore, 1997). Learner-instructor interactions are generally viewed as the most important as they entail the interactions that occur between student and teacher though they may be quite varied in how they occur, such as through synchronous videoconference, asynchronous emails, or other (Duc, 2012; Elyakim et al., 2019; Grigoryan, 2017; Moore, 1997). Finally, knowledge sharing through peer discussions between students is the highlight of learner-learner interactions (Boyle et al., 2010; Duc, 2012; Elyakim et al., 2019; Larkin & Jamieson-Proctor, 2015; Yilmaz, 2017).

While learner-learner, learner-content, and learner-instructor interactions dominate the definition of dialogue, other interaction combinations have also been suggested. Hillman et al. (1994) advised adding learner-interface. In a f2f environment, learner-interface interactions would include voice inflections, body language, and vocabulary, while online learner-interface interactions would include any of the technologies used to communicate or the use of the online learning materials (Dockter, 2016; Hillman et al., 1994; Newkirk et al., 2013; Weidlich & Bastiaens, 2018). While not fully accepted by all TDT researchers, some have begun to embrace the fourth interaction of learner-interface (also known as learner-technology and learner-technology interface) as a meaningful component of understanding dialogue (Elyakim et al., 2019; Ismail et al., 2018; Quong et al., 2018; Weidlich & Bastiaens, 2018; Yilmaz, 2017; Yilmaz & Keser, 2017). Additionally, Shin (2002) made the argument for adding learner-institution interactions, and Menchaca and Bekele (2008) have recommended that measures of success be included. While these interactions and others have been proposed, further research into their actual effect on dialogic transactional distance still needs to be conducted (Ekwunife-Orakwue & Teng, 2014).

The effect of dialogue on transactional distance can be somewhat viewed as an inverse relationship where the more dialogue there is, the less transactional distance. Moore (1997) suggested that the quantity of dialogue did not matter, but that quality dialogue was what had the greatest effect on reducing teacher-learner psychological mismatch. The extent and nature of the dialogue can be influenced by many factors including the difficulty of the course subject, teacher and student personalities, ease of use and design of the course, teacher and student physical environments, the number of students a teacher must have the dialogue with, or teacher and student support, to name a few (Moore, 1997; Newkirk et al., 2013; Weiss & Belland, 2018). Each of these factors can play a part in whether a teacher and student can have a quality conversation without outside distractions or influences. For example, if an instructor must only communicate with five students, then the quality of those dialogues can be much higher than if the instructor had to communicate with 30.

Moore (1997) proposed that, generally, asynchronous interactions often found in distance education courses created a slowed dialogue and increased transactional distance over synchronous interactions. The medium used, whether it be asynchronous chat boards or synchronous videoconferencing, is not enough to lower transactional distance (Moore, 1997). This may be because teachers may not use the medium as intended or to its fullest potential, or students may prefer not to respond or interact. Moore suggested that dialogue is often higher in upper-level and college courses, in smaller group settings, when Socratic teaching approaches are used, and in the social sciences. Dialogue in more teacher-centered courses, such as science and mathematics, is often much lower, giving rise to higher transactional distance (Moore, 1997).

With the early definitions and lack of communication media, transactional distances were often considered quite high (Krieger, 2017; Moore, 1997). The evolution of communicative technologies decreased the amount of separation between students and the many aspects of distance learning but forced an expanded definition of dialogue to include all the possible relational interplays at work in a distance education environment (Krieger, 2017; Moore, 1997; Shin, 2002). As the definition of dialogue had an overhaul over the years, so too has the definition of structure.

Transactional Distance Theory – Structure

Program structure, like dialogue, is thought to reduce transactional distance when there is more of a give-and-take relationship between the student and the subject. Original definitions of structure described the various means through which study material could be presented, student work could be assessed, and the flexibility of the course to be adapted to students' needs (Moore, 1997). In the initial stages, structure mostly referred to post-sent materials and presented a rigid, predetermined format, or video/teleconferencing which presented a flexible structure that could more easily accommodate student needs (Moore, 1997). It was suggested that transactional distance could be lowered if the course content and material were interactive enough that the student had some voice over course organization, implementation of materials, and assessment procedures (Krieger, 2017; Moore, 1997; Quong et al., 2018; Yilmaz & Keser, 2017).

Some researchers have proposed that structure and dialogue may be inversely related (e.g., Larkin & Jamieson-Proctor, 2015; Moore, 1989, 1997; Quong et al., 2018; Saba & Shearer, 1994). This would mean that when a course is highly structured, there is little room for teacher-student dialogue, student-course interaction, or course adaptations resulting in high transactional distance between teacher and student (Moore, 1989, 1997; Saba & Shearer, 1994). If the course structure is less controlled it allows for more student-teacher dialogue and can more easily be adapted to student needs lowering transactional distance (Moore, 1989, 1997; Saba & Shearer, 1994). More recent terminology for this would be individualized, personalized, and differentiated learning.

Newkirk et al. (2013) suggested that course structure should be thought of as being comprised of two components, course organization and course delivery. Moore (1997) pointed out that creating an environment in which course delivery is interactive with the student while also organizing the course to maintain appropriate rigor, pacing, sequencing, and appropriate opportunities for student practice and feedback is often extremely difficult to achieve and can often require the skills of several individuals working together to develop. According to Moore (1997), six elements should be addressed to have a successful distance education program:

- The presentation of materials must be appropriately sequenced and offered in a format best suited to the information. For example, demonstrations would be best presented by videos, while diagrams should be in a text format.
- Course materials and structure should be created in such a way as to stimulate students' interest and maintain student motivation.
- Instructors must design courses and materials to promote students' higher-order critical thinking skills, such as analysis and critique. Moore (1997) admits this may be the most difficult element to achieve in distance education courses.
- Students must be provided with guidance on how to study materials, and there must be a way for students to ask questions and have them answered in a timely manner.
- Students must be provided with opportunities to practice the skills they are learning and be evaluated on their knowledge and application of the material.
- Students must be provided with opportunities to share and discuss what they are learning with others to create new understandings and knowledge.

Again, as with dialogue, the definition of structure matured with the advancement of available technology. Over the years, researchers have conducted many studies to attempt to pinpoint the elements of structure that help to lower transactional distance (e.g., Benson & Samarawickrema, 2009; Chen & Willits, 1999; Huang et al., 2015;

Ismail et al., 2018; Kearsley & Lynch, 1996; Krieger, 2017; Larkin & Jamieson-Proctor, 2015; Mbwesa, 2014; Newkirk et al., 2013; Stein et al., 2005; Yilmaz, 2017; Yilmaz, & Keser, 2017). Some of these elements include student satisfaction with the perceived knowledge gained (Krieger, 2017; Larkin & Jamieson-Proctor, 2015; Stein et al., 2005), student satisfaction with course setup (Krieger, 2017; Larkin & Jamieson-Proctor, 2015; Mbwesa, 2014; Yilmaz & Keser, 2017), and the flexibility of course organization and delivery including the use of social media and videoconferencing (Chen & Willits, 1999; Elyakim et al., 2019; Ismail et al., 2018; Krieger, 2017; Yilmaz, 2017). As technology has evolved, the definition of this component of TDT has taken shape to include structure mainly as it pertains to learner-content and learner-interface interactions (Benson & Samarawickrema, 2009; Huang et al., 2015; Krieger, 2017).

Structure and dialogue were originally the only two elements included in the transactional distance theory. However, it was not long before Moore (1989, 1997) realized that there was a third element at play – one that the instructor could have little control over – the ability of the learner to work and learn independently.

Transactional Distance Theory – Learner Autonomy

Stereotypical images of brick-and-mortar school formats traditionally find one envisioning an instructor-centered classroom in which students sit at desks in neatly lined rows facing the teacher, who most likely has his or her back to them as he/she relentlessly scribbles notes on a chalkboard or whiteboard for students to copy. These students would be the opposite of what Moore (1989, 1997) had in mind of an autonomous learner. Traditional school formats train students to be dependent upon an instructor for information and guidance instead of training students on how to take control of their own learning (Kontkanen et al., 2017; Moore, 1997; Swart & MacLeod, 2021; Vickrey et al., 2018; Weiss & Belland, 2018; Yilmaz, 2017). The ideal autonomous learner does not even need a teacher but instead is self-directed, has a high internal motivation, is capable of locating and utilizing resources for their own learning, and is satisfied with the prospect of getting to learn as a motivator rather than needing rewards or outside recognition from superiors (Elyakim et al., 2019; Larkin & Jamieson-Proctor, 2015; Moore, 1997; Quong et al., 2018). As Duckworth and Gross (2014) put it, the ideal autonomous learner has the grit to be in complete control of his or her own learning. Because an ideal learner is extremely rare, most researchers agree that learner autonomy is the extent to which the learner takes control of his or her learning by influencing the content, pacing, evaluation methods, and educational goals for their personal learning experiences (e.g., Elyakim et al., 2019; Krieger, 2017; Larkin & Jamieson-Proctor, 2015; Moore, 1997;).

Krieger (2017) suggested that, unlike dialogue and structure, there has been little controversy over a definition of learner autonomy, perhaps due to the lack of educator control over the element. “Instructors cannot drive internal motivation, no matter the type of student” (Krieger, 2017, p. 17). However, Macaskill and Taylor (2010) argued that the definition of an autonomous learner has varied from being the student’s ability to learn in a format or manner that he or she chooses, the student’s ability to decide what he or she learns, or the student’s initiative in being persistent and resourceful with his or her learning. Macaskill and Taylor’s (2010) discontent with the varied definitions of an autonomous learner is what led to their development of the Autonomous Learning Scale. This scale has created the most recent and agreed-upon definition of an autonomous learner to date. New aspects of the definition include:

> that autonomous learners take responsibility for their own learning, are motivated to learn, gain enjoyment from their learning, are open-minded, manage their time well, plan effectively, meet deadlines, are happy to work on their own, display perseverance when encountering difficulties and are low in procrastination when it comes to their work. (Macaskill & Taylor, 2010, p. 357).

Limitations and Assumptions of Transactional Distance Theory

Over the years, Moore's (1997) theory has been met with some critique, particularly in the definitions of dialogue, structure, and learner autonomy (e.g., Gorsky & Caspi, 2005; Krieger, 2017; Larkin & Jamieson-Proctor, 2015; Papadopoulos & Dagdilelis, 2006). The lack of true operational definitions for the three elements made researching the relationships between them quite difficult. This led to Gorsky and Caspi's skepticism of the theory in the early 2000s (Gorsky & Caspi, 2005; Krieger, 2017; Larkin & Jamieson-Proctor, 2015; Weidlich & Bastiaens, 2018; Weiss & Belland, 2018). As technology has progressed, so too have the definitions, extending the capabilities for more valid research to be conducted and reducing the significance of Gorsky and Caspi's (2005) criticism (Krieger, 2017). However, Huang et al. (2015) suggested that while TDT has great value to pedagogy, its value to education can be amplified with more concrete definitions and empirical research that assesses the relationships between the elements. The issue most concerning to researchers is that as technology continues to evolve, so too must the definitions of dialogue, structure, and learner autonomy (e.g., Huang et al., 2015; Krieger, 2017). This creates a vicious cycle of a call for finite definitions and an obligation for the definitions to advance with the technology leaving TDT in a continuous custody battle between researchers who claim it is not valuable and those who assure it is (e.g., Huang et al., 2015; Krieger, 2017; Moore, 1997; Papadopoulos & Dagdilelis, 2006; Weiss & Belland, 2018).

Transactional Distance Theory in Relatable Studies

As noted earlier, TDT has been one of the most prominent theories for distance education (Duc, 2012; Elyakim et al., 2019; Garrison, 2000; Grigoryan, 2017; Ismail et al., 2018; Krieger, 2017; Larkin & Jamieson-Proctor, 2015; Murphy & Rodriguez-Manzanares, 2008; Quong et al., 2018; Weidlich & Bastiaens, 2018; Yilmaz, 2017; Yilmaz & Keser, 2017). As such, it has been utilized as the leading theory in many distance education research studies including those by Chen and Willits (1999); Grigoryan (2017); Ismail et al. (2018); Larkin and Jamieson-Proctor (2015); Quong et al. (2018); Saba and Shearer (1994); Swart and MacLeod (2021); Weidlich and Bastiaens (2018); Yilmaz (2017); and Yilmaz and Keser (2017). Unfortunately, each of these studies focused on university students or the more advanced graduate students. Few studies seemed to consider high school students and the transactional distances they faced in distance education courses. Those that did were generally much older than those involving university students. Only Elyakim et al.'s (2019) and Weiss and Belland's (2018) studies included both high school students and TDT in the last 5 years. The studies on university students and high school students tended to come to similar conclusions…that frequent, purposeful communication between students and instructors is one of the most impactful aspects of reducing transactional distance and increasing student success in online courses (Dockter, 2016; Garthwait, 2014; Velasquez et al., 2013; Yilmaz, 2017; Yilmaz & Keser, 2017), but that the technology, if not used ideally, could impede progress (Elyakim et al., 2019; Ismail et al., 2018).

In a study of university students taking an online course, TD was found to be lower in students who took the course asynchronously in discussion forums versus synchronously with web-conferencing, and lower in students who were provided with metacognitive prompts to support student's awareness of their learning processes (e.g., Yilmaz & Keser, 2017). Additionally, significant differences were found in student achievement when the learning environment consisted of asynchronous discussions that involved metacognitive support versus synchronous courses with the same support (e.g., Yilmaz & Keser, 2017). Similarly, Yilmaz (2017) found that in asynchronous environments, such as Facebook, students were more engaged in knowledge sharing when there was lower perceived TD, and that TD played a greater part in promoting knowledge sharing behaviors than perceptions of social presence.

In their study of perceived TD amongst online instructional faculty, Murphy and Rodriguez-Manzanares (2008) found that despite the use of advanced technologies, a strict curriculum, inflexible learning platform, and synchronous voice or videoconferencing could all be detrimental to the goal of lowering TD and increasing students' online performance. Additionally, both Papadopoulos and Dagdilelis (2006) and Quong et al. (2018) found that the learning platforms, social media platforms, or computer programs used could cause higher transactional distance, even in f2f courses, due to the lack of understanding and "dislike" of the technology. Findings from both studies suggested that while technology could be used to lower transactional distance by increasing the possibility of improved dialogue and content understanding, unless students fully comprehended how to use that technology, or if the technology distracted from learning (Quong et al., 2018), transactional distance will often increase (Papadopoulos & Dagdilelis, 2006). Furthermore, Ismail et al. (2018), suggested that some of the challenges of using e-learning tools were that learners may feel isolated which can negatively affect their academic progress.

Swart and MacLeod (2021) also found that technology, even in problem-based learning classrooms, was not enough to lower TD. In their study, classrooms that were designed to incorporate more technology and promote problem-based learning did not significantly improve student learning, progress, or satisfaction versus a normal, nontiered, classroom in which students were allowed to rearrange desks to face each other (tiered classrooms had much lower results on the previously stated criteria and had higher perceived TD). The students in the problem-based learning classrooms and the flat classrooms were better able to have discussions with each other as students were able to sit f2f. In the tiered learning classrooms, however, students were still able to sit somewhat in groups, though usually with some member sitting in a row in front of the rest of the group and having to turn around awkwardly to talk to them. Swart and MacLeod's (2021) study suggested that even in f2f learning groups where physical location is not a factor, perceived TD can vary based on something as simple as seating configurations.

In nonf2f high school courses, Elyakim et al. (2019), found that TD could be lowered when students were given the opportunity to use location-based mobile learning as opposed to class discussions. The location-based mobile learning context provided students with recommended sequence suggestions, location assignments, navigational tools, and instructional information while on a class field trip (Elyakim et al., 2019). The location-based mobile learning was found to lower TD because, while there was still a physical distance from a teacher and the course was entirely online, the location in which the students engaged in their learning became part of the learning material suggesting that technology alone could not lower TD, but that the context in which learning occurred could (Elyakim et al., 2019).

Weiss and Belland's (2018) case study examined how problem-based learning groups in high school science classes can function autonomously when supported by structured, computer-based scaffolding. Weiss and Belland (2018) suggested that learning units structured to "(1) employ a simplified problem-solving framework, (2) provide repeated use of the framework, increasing both familiarity and confidence using the approach, and (3) frame the use of dialog to encourage student/student and student/interface dialog" (p. 98) meet the criteria for decreasing TD and improving learning group autonomy, rather than just individual learner autonomy. Additionally, Swart and MacLeod (2021) found that within various f2f settings, TD could be lowered.

Velasquez et al. (2013) considered how technologies used in online courses facilitated students' perceptions of being cared for, thus promoting student progress and success. Velasquez et al. (2013) suggested that TD could be decreased as students felt a greater amount of caring from their teachers, best achieved with synchronous conversations. However, asynchronous conversations may also lead to lower transactional distance and more perceived caring so long as the responses were prompt, reflective, and of high quality (Velasquez et al., 2013). Comparatively, Dockter (2016) argued that TD is immediately increased when communication was through writing as every text was interpreted differently by students due to their unique backgrounds. Dockter (2016) suggested TD could be lowered when educators "presented" themselves to students through frequent, regular, and varied opportunities for communication. Additionally, Grigoryan (2017), found that the use of audiovisual commentary as feedback on student assignments increased student satisfaction and perceived understanding of course material, while the personalized interaction lowered TD.

Rational for Transactional Distance Theory Over Other Frameworks

TDT, originally proposed 50 years ago, is considered one of the founding theories for distance education (Dockter, 2016; Elyakim et al., 2019; Gorsky & Caspi, 2005; Grigoryan, 2017; Krieger, 2017; Larkin & Jamieson-Proctor, 2015; Quong et al., 2018; Swart & MacLeod, 2021; Weidlich & Bastiaens, 2018; Yilmaz, 2017; Yilmaz & Keser, 2017). As such, it has stood the test of time and scrutiny and remains one of the most mentioned and utilized theories in distance education research. More specifically, TDT provides a means of considering the psychological differences between students and teachers experienced in both an online environment and a f2f environment allowing for more comprehensive discussions regarding the comparability of the two environments to be made.

While most of the research based on TDT has been focused on the university level student (e.g., Grigoryan, 2017; Larkin & Jamieson-Proctor, 2015; Quong et al., 2018; Swart & MacLeod, 2021; Weidlich & Bastiaens, 2018; Yilmaz, 2017; Yilmaz & Keser, 2017), several researchers have begun to tie its importance to the high school level student and the educational experience had (e.g., Elyakim et al., 2019; Garthwait, 2014; Murphy & Rodriguez-Manzanares, 2008; Papadopoulos & Dagdilelis, 2006; Velasquez et al., 2013; Weiss & Belland, 2018). This is not to say, however, that other frameworks or theories could have sufficed as a lens through which to view this research. Two other major frameworks considered were Koehler and Mishra’s (2005) theory of the interplay between teachers’ technological, pedagogical, and content knowledge, known as TPACK, and Kearsley and Schneiderman's (1998) engagement theory.

Originally proposed as a framework for what teachers need to know to be effective, Shulman (1986, 1987) suggested that there must be a blending between content knowledge and pedagogical knowledge when designing learning activities. Over the years, as technology became increasingly advanced, Koehler and Mishra (2005) added a technological element. Koehler and Mishra’s (2005) technological, pedagogical, and content knowledge (TPACK) framework was designed to help teachers be more effective in the digital age (Atun & Usta, 2019; Huang, 2018). One of the greatest advantages of TPACK is that, as a general framework, it is not tied to a specific technology or topic

(Atun & Usta, 2019; Kontkanen et al., 2017). Due to its flexibility, TPACK has been adapted over recent years to be subject-specific, for example, technology-pedagogy-science-knowledge (TPASK) or technology-pedagogy-mathematical-knowledge (TPMK), while remaining a relevant and valuable concept (Huang, 2018).

The TPACK framework, as a concept, has years of backing; however, it has fallen short of expectations when put into practice. Although Koehler and Mishra's (2005) TPACK framework is a highly accepted framework, there are aspects of the theory that make it unsuitable for this study. First, as Huang (2018), Kontkanen et al. (2017), and Matherson et al. (2014) pointed out, the framework sounds great as a concept but has yet to unveil itself as a practical framework. Second, the TPACK framework is a complex coalescence of several domains with no clear boundaries between them making it "difficult to create instruments for measuring and assessing TPACK domain" (Kaplon-Schilis & Lyublinskaya, 2020, p. 26). Third, the TPACK framework is intended to be used by teachers as a model of how to converge multiple knowledges to become a more successful instructor; to be used by researchers to consider how well technology is being incorporated; and by teachers, supervisors, and researchers alike to determine whether instructional techniques are effective (Atun & Usta, 2019; Dikmen & Demirer, 2022; Kaplon-Schilis & Lyublinskaya, 2020). TPACK was not intended to be used as a tool for comparing learning environments regarding students' success, at least not in the way as is intended by this study. Perhaps future studies could consider using the TPACK framework to evaluate the effectiveness of the instructional design of the two environments, but that was beyond the scope of this study.

Kearsley and Schneiderman's (1998) engagement theory is essentially a theory designed to describe methods used to promote student interactions with learning content and others through meaningful learning activities and worthwhile tasks. Although engaged learning can occur in a f2f classroom, Kearsley and Schneiderman (1998) intended the theory to be used as a framework for technology-based teaching and learning.

Successful engaged-learning environments include some aspects of human interaction and should be student-centered (Francis, 2018; Hoskins, 2012; Knowlton, 2000). Having human or social interaction in learning environments promotes active learning from students while also producing higher student satisfaction, learning achievement, and student retention (Elyakim et al., 2019; Hoskins, 2012; Ismail et al., 2018; Swart & MacLeod, 2021; Yilmaz & Keser, 2017). Knowlton (2000), however, argued that online course activities alone require students to be active, engaged learners by simply requiring them to click on and open a hyperlink within their course, or to post to a discussion to make their presence known without necessarily having to have human interaction.

Reviewing Knowlton's (2000) argument along with Francis's (2018) review of engagement theory leads to some uncertainty as to what exactly constitutes student engagement. Francis (2018) suggested that there were varied meanings of engagement across levels of learning, participant perspectives, and ideologies, and Knowlton (2000) made a clear case as to the varied thoughts as to how little student involvement must take place to be considered as engaged learning, i.e., clicking a hyperlink. The lack of a unified definition of engagement is one of this theory's major disadvantages. Although nearly 2 decades old, engagement theory is still very young and new, and has yet to be "put through the wringer" in having its definition's varying aspects truly scrutinized and pulled together comprehensively as to be useful (Francis, 2018), least to this study.

Theories, in general, are a way of making sense of the research conducted in any field. Learning theories are a way to understand how students learn, and what steps can be taken to improve and enhance the learning experience (Learning Theories and Student Engagement, 2014). Engagement theory in its simplest form, like so many other learning theories dominating the field, specified that students must be actively involved in their learning (Francis, 2018; Hoskins, 2012; Irwin et al., 2014; Knowlton, 2000; Learning Theories and Student Engagement, 2014).

TDT also suggests that students be involved in their own learning but takes engagement a step further by identifying that students be actively engaged in conversations with their peers and instructors, that students who are actively engaged are autonomous learners, and that the design of the course plays a role in how engaged students can be (Hoskins, 2012; Learning Theories and Student Engagement, 2014). Unfortunately, most educational theories, such as engagement theory, TPACK, constructivism, or transformative learning, do not allow for the in-depth comparison investigation of the online and f2f environments that TDT does because they all lack the ability to consider the relationships between teacher-student, student-student, teacher/student-environment, or the student's ability to be an independent learner. TDT has been widely accepted and adopted across many learning platforms and levels (e.g., Duc, 2012; Garrison, 2000; Krieger, 2017; Murphy & Rodriguez-Manzanares, 2008; Quong et al., 2018; Swart & MacLeod, 2021; Weiss & Belland, 2018), and lends itself as being of greater worth to this study than other considered frameworks.

As noted earlier, much of the research on TDT is either based on postsecondary education or, if the focus is regarding the high school level, is out-of-date with today's technological advances. Even within the past 5 years, the level of understanding of the need for online learning has blossomed tremendously. However, the lack of recent research concerning TDT in secondary education made it near impossible to base this study on current understandings. One of the goals of this study was to aid in updating the literature in this field. The following sections are intended to draw parallels between high school Algebra I and the three tenets of TDT. Where possible, recent research was used to validate these associations. Where recent research was not available, older, more relevant references were used.

High School Algebra I

High School Algebra I and Dialogue

The tenet of dialogue in transactional distance theory refers to the positive, two-way interactions between students and teachers, students and students, students and content, or students and interface (Elyakim et al., 2019; Grigoryan, 2017; Larkin & Jamieson-Proctor, 2015; Moore, 1997; Quong et al., 2018). Research suggested that whether the dialogue is asynchronous or synchronous can play a part in both the quantity and quality of dialogue that can occur (e.g., Dockter, 2016; Huang et al., 2015; Moore, 1997; Protopsaltis & Baum, 2019; Weiss & Belland, 2018). Additionally, the quantity and quality of the dialogue that takes place within an educational environment can have a great impact on student engagement (e.g., Dumford & Miller, 2018; Moore, 1997; Protopsaltis & Baum, 2019; Quong et al., 2018; Weiss & Belland, 2018; Wheatley, 2016), student satisfaction (e.g., Protopsaltis & Baum, 2019; Quong et al., 2018; Weidlich & Bastiaens, 2018), student motivation (Wheatley, 2016), student perceived learning (Quong et al., 2018), and on the probability of students persisting through a course to completion (e.g., Heinrich et al., 2019). It is also believed that dialogue can provide a humanistic feeling of community to online students who may otherwise feel isolated (Dockter, 2016; Ismail et al., 2018; Moore, 1993). However, dependent on the course content, content level, and class size, dialogue may not be as high quality and could potentially lead to greater feelings of dissatisfaction and loneliness (e.g., Dockter, 2016; Heinrich et al., 2019; Moore, 1997; Usta & Mirasyedioğlu, 2022). Dialogue may also be of lower quality and quantity in less motivated students (Ezen-Can & Boyer, 2015), in lower-level courses, teacher-centered courses, or the natural sciences, such as mathematics and science, further contributing to feelings of dissatisfaction and loneliness (Moore, 1997; Protopsaltis & Baum, 2019).

Moore (1997) originally proposed that online asynchronous dialogue meant that communication was slower than what it would be during synchronous communications leading to higher transactional distance. Protopsaltis and Baum (2019) suggested that online synchronous communication tools can be great for improving dialogue. However, Ismail et al. (2018) wrote “students in an online learning environment lack opportunities to experience the benefits of both structured dialogue and the sense of community that can be created in the more traditional classroom environment" (p. 5), and some research suggested that online tools that allowed synchronous video- or teleconferencing often made it difficult to control multiple users trying to talk at the same time (e.g., Murphy & Rodriguez-Manzanares, 2008; Yates et al., 2021). Synchronous video or teleconferencing tools could create nearly inaudible conversations and often forced teachers to mute users, resulting in teacher-dominated conversations and diminished opportunities for student-to-student collaborative learning (e.g., Murphy & Rodriguez-Manzanares, 2008; Yates et al., 2021). Regardless of these findings, several researchers continued to state that online asynchronous communication could be as effective as f2f synchronous communications in lowering transactional distance if teachers supplied prompt, reflective responses and attempted to create responses that are in tune with student personalities (e.g., Kontorovich, 2018; Larkin & Jamieson-Proctor, 2015; Velasquez et al., 2013; Yilmaz & Keser, 2017).

Some researchers argued that f2f synchronous communications were better suited for some lower-level or high school students because these students may not have had the willingness to ask for help or otherwise communicate online (e.g., Protopsaltis & Baum, 2019; Viano, 2018; Yates et al., 2021), and it may have been easier for an instructor to adapt the material to adjust for student learning and understanding in the f2f environment (e.g., Dockter, 2016; Moore, 1997; Yates et al., 2021). The general perception of the successful online student tended to be that of a highly motivated, autonomous learner with good time management and communication skills, or in other words, a university-level student (Protopsaltis & Baum, 2019; Yates et al., 2021). However, some research suggested that student age and prior online experience did not affect dialogic interactions and, in turn, had no effect on student achievement (e.g., Ekwunife-Orakwue & Teng, 2014). Additionally, some research indicated that students were more engaged and generally believed in having learned more in their online courses (e.g., Elyakim et al., 2019), especially if they believed their teachers care about them (e.g., Barnett, 2016; Velasquez et al., 2013), and they believed their teachers were engaged in their learning (e.g., Frazelle, 2016; Grigoryan, 2017; Velasquez et al., 2013; Yates et al., 2021). Furthermore, students were more satisfied in their online courses when learning occurred in interactive environments that allowed for the construction and sharing of knowledge (Grigoryan, 2017; Yilmaz, 2017), such as when using a learning platform that resembled frequently used social networking sites (for example, Facebook or Twitter) and allowed for personalization (e.g., Quong et al., 2018; Yilmaz, 2017). These studies make an argument for online learning to be just as effective for high school learning as it would be for university-level students if certain criteria were met.

Unfortunately, assumptions that can be made for average to advanced high school students may not hold as well for at-risk ninth-grade Algebra 1 students. Research suggests younger students and students who were at risk to fail in general, were more likely to have significantly lower technology skills or have skill deficits, such as reading and writing, which were not necessarily tied to the failed course but may prohibit a student’s ability to communicate in an asynchronous discussion (e.g., Heinrich et al., 2019; Viano, 2018).

Dialogue in online secondary schools is especially important as high school students tend to need more assistance in online classes to better deal with the struggles of isolation and lack of technical skills (e.g., Protopsaltis & Baum, 2019; Viano, 2018; Wheatley, 2016), or to better help students truly learn the material, especially those who rush through assignments without fully understanding it (e.g., Belland et al., 2019). Unfortunately, neither high school nor university-level studies have reported dialogue to be successfully linked to predicting student achievement (e.g., Ekwunife-Orakwue & Teng, 2014), or course satisfaction (e.g., Weidlich & Bastiaens, 2018). Course satisfaction has been linked with both academic engagement and academic achievement (Abuhassna et al., 2020; Buzzai et al., 2021; Gopal et al., 2021).

The studies that most closely linked dialogue and academic achievement were those by Kontorovich (2018), Libasin et al. (2021), Muhonen et al. (2018), and Yilmaz and Keser (2017). Kontorovich (2018) reported that students in a f2f high school Algebra class were more likely to score higher on assessments and have higher course grades if they were more engaged in an online discussion forum with their classmates. Libasin et al. (2021) compared the effect of synchronous and asynchronous communications on student achievement for university-level students in math classes during the early stages of the COVID-19 pandemic. The results of Libasin et al.'s (2021) study found that students in the synchronous courses (via Google Meet) were more likely to pass the courses than students who took the asynchronous courses. Muhonen et al. (2018) found that quality dialogue, either teacher-initiated or student-initiated, resulted in higher grades for sixth-grade language arts and physics/chemistry classes, but no significant differences were found in religion, history, or biology/geography classes. However, this study does not take into account teacher differences between the courses taught, does not employ factors other than class size and teacher qualification to determine the effect on student achievement, and does not consider measures of student achievement other than the course grades that were given by the teachers participating in the study. Yilmaz and Keser (2017) found significant differences in undergraduate student achievement when metacognitive support was combined with asynchronous forms of communication over synchronous web conferences with the same support. However, the main considerations of this study were the medium used for dialog, not necessarily the quantity or quality. Still, some have clung to the belief that online learning has the potential to provide flexible, self-paced learning environments that motivate and engage students on an array of topics, including Algebra I and credit recovery (e.g., Heppen et al., 2017; Viano, 2018).

According to Protopsaltis and Baum (2019), many educators were “skeptical about the quality and online learning and its pedagogical value” (p. 4). Some of the greatest concerns reported by educators when it comes to online learning included students’ technology skills (Abuhassna et al., 2020), “course quality, academic integrity, and lack of f2f interaction” (Clements, Zweig, et al., 2015, p. 2). When investigating teacher perceptions of transactional distance, several researchers found that most teachers were more satisfied in a traditional f2f course where they were able to interact more closely with students and felt there was a greater transactional distance online due to less dialogue and larger numbers of students (Usta & Mirasyedioğlu, 2022; Wengrowicz &

Offir, 2013). Lu (2011) agreed with these teachers by stating that “communication is one of the most difficult problems instructors may have to deal with when teaching online courses, especially mathematical courses” (p. 1). It was also suggested “that only one third of the students contribute to the [online] discussions, while another one third of them only read what is written, and the rest never logged in to the online environment” (Yilmaz, 2017, p. 845). Kontorovich (2018) also reported that only a small number of students accounted for the majority of the online discussions.

Regardless of the above, some researchers argued that no matter the course communication and dialogue could be increased by combining asynchronous and synchronous tools (Protopsaltis & Baum, 2019; Yilmaz & Keser, 2017), using medium other than text to communicate (Dockter, 2016), promoting location-based mobile learning where the location becomes part of the learning material (Elyakim et al., 2019), providing feedback that includes both audiovisual components as well as text (Grigoryan, 2017) and making online courses smaller in the number of students (Dumford & Miller, 2018). Additionally, some studies indicated that when students “lurked” or viewed what other students and instructors had posted before doing their own writing then they were more likely to become actively engaged in their learning by creating posts themselves or responding to others’ posts (Quong et al., 2018).

High School Algebra I and Structure

When it comes to TDT’s concept of structure, the general perceptions of distance education researchers suggested that structure is the overall organization of the course and the delivery methods used to disperse and collect learning material (Elyakim et al., 2019; Krieger, 2017; Larkin & Jamieson-Proctor, 2015; Moore, 1997; Quong et al., 2018; Yilmaz & Keser, 2017). Structure elements are mostly considered on a spectrum of flexibility where the more flexibility there is, the more differentiation and personalization can be made, potentially lowering the transactional distance (Elyakim et al., 2019; Moore, 1997). These elements could include the difficulty of the learning platform, the teaching strategies and evaluation methods utilized, or the emotional characteristics of both the teachers and students (Dockter, 2016; Moore, 1997; Yilmaz, 2017; Yilmaz & Keser, 2017). Within the aspect of structure, there are two types of interactions that can occur, learner-content interactions or learner-technology, also called learner-interface interactions.

According to Quong et al. (2018), learner-content interactions are one of the main purposes of education, to learn the material. These interactions are often influenced by the formality and flexibility of the course subject and course design (Elyakim et al., 2019; Quong et al., 2018). For example, subjects that require more regulated sequencing, such as most mathematics courses, do not allow for as much flexibility as subjects that do not have specific order constraints, such as many of the social sciences and art classes (Garrison & Cleveland-Innes, 2005; Moore, 1997; Usta & Mirasyedioğlu, 2022).

Learner-interface interactions include the ability of the participants to navigate and use online tools (Quong et al., 2018), including the tools used to communicate. These interactions were often dependent on the choice of media used, including the instructor's and student's knowledge of how to use it, the functionality, visualization, and overall usability of the media and course tools (Quong et al., 2018).

According to Goel et al. (2012), "students need an interactive constructive environment to allow for construction of knowledge" (p. 12), which can be achieved through a properly designed online course. Unfortunately, creating and maintaining online courses that use appropriate tools and are suitably designed while maintaining flexibility are very difficult to do (Koehler & Mishra, 2005; Moore, 1997; Protopsaltis & Baum, 2019; Usta & Mirasyedioğlu, 2022), and often result in many online teachers simply transferring f2f teaching methods and materials to the online environment, losing potential engagement factors for students (Dumford & Miller, 2018; Huang, 2018; Kontkanen et al., 2017; Vickrey et al., 2018).

Even if an online course was designed to allow for maximum flexibility while maintaining appropriate structure, does not mean students were able to succeed. Students in online courses must have the necessary skills to operate online course management systems (Papadopoulos & Dagdilelis, 2006; Quong et al., 2018), including an understanding of the technology and the technology languages in order to be successful (Papadopoulos & Dagdilelis, 2006; Wakil et al., 2019). When students did not have the necessary skills to adequately navigate the online environment, such as reading proficiency or knowledge of how to find items in a menu bar, they tended to spend more time trying to learn the technology than they did trying to learn the material (e.g., Heinrich et al., 2019; Papadopoulos & Dagdilelis, 2006; Quong et al., 2018). Some difficulties may be that the way in which the environment operates was counter-intuitive to what the learner expects (Papadopoulos & Dagdilelis, 2006), or, in the case of mathematics, the learning platform or environment did not allow for necessary symbols to be used (Usta & Mirasyedioğlu, 2022) making the learning of material even more difficult. Additionally, Dumford and Miller (2018) found that university students who took Abstract Algebra online spent more time reviewing lessons and requesting f2f hours than students who took the same course entirely f2f and that the online students did not perform as well as the f2f students on the Abstract Algebra problem-solving test. Dumford and Miller's (2018) findings combined with those of Libasin et al., (2021), Usta and Mirasyedioğlu (2022) and Heppen et al. (2017) suggested that teaching and learning mathematics online was perceived as more difficult than in the f2f counterparts, often due to the structure of the online learning platforms, increasing transactional distances.

Even in online Algebra I courses, however, there was potential for success. While it was suggested that students tended to separate personal and academic uses of technology (Quong et al., 2018), several researchers have suggested that students preferred online courses that were structured to resemble popular social media sites that they were familiar with, such as Twitter and Facebook (Quong et al., 2018; Yates et al., 2021; Yilmaz, 2017). If students were more familiar with the overall structure of the course, then they may have found them easier to use (Quong et al., 2018; Yilmaz, 2017). Prior experience, such as when a course resembles familiar websites, combined with location-based mobile learning (Elyakim et al., 2019), could be enough to influence a student's course satisfaction and engagement to lower transactional distance and improve academic achievement.

Prior experience, computer training, and technology self-efficacy have yet to be able to predict student success or course performance (DeTure, 2004; Waschull, 2005), however, learner-content interactions may be more significant in predicting student achievement than learner-technology interactions (Ekwunife-Orakwue & Teng, 2014; Elyakim et al., 2019; Kim et al., 2021). Additionally, the richness of the media used in an online course has been found to have an impact on learner-technology and learner-content interactions which were believed to have about twice as much impact on student satisfaction than learner-instructor or learner-learner interactions (Ekwunife-Orakwue & Teng, 2014).

Moore (1997) suggested that the more flexibility there was in a course design, the more potential for lowered transactional distance, especially if the structure was designed with deliberate activities to build rapport, increase collaboration, and increase dialogue with active, accessible teachers (e.g., Elyakim et al., 2019; Forte et al., 2016; Murphy & Rodriguez-Manzanares, 2008; Velasquez et al., 2013; Yates et al., 2021). However, the greatest problems in achieving effective structures such as this were when teachers used the online platform in a way more suitable for a f2f course (Huang, 2018; Kontkanen et al., 2017; Usta & Mirasyedioğlu, 2022; Vickrey et al., 2018), when the learning platform was not liked, unfamiliar, or lacked usability (e.g., Murphy & Rodriguez-Manzanares, 2008; Quong et al., 2018; Yilmaz, 2017), or when there was an imposed, strict curriculum which restricted the amount of flexibility of the structure and content (e.g., Elyakim et al., 2019; Murphy & Rodriguez-Manzanares, 2008; Usta & Mirasyedioğlu, 2022).

High School Algebra I and Learner Autonomy

The final element of TDT is that of learner autonomy. Learner autonomy refers to the learners' ability to determine their own learning goals, course pacing and evaluation procedures, and their ability to be self-directed, active participants who do not need to be constantly monitored or rewarded by authoritative figures (Dockter, 2016; Larkin & Jamieson-Proctor, 2015; Moore, 1997; Yates et al., 2021). According to Moore (1997), ideal autonomous learners are those individuals who can completely teach themselves, and that autonomous learners, in general, tend to prefer less dialogue and less structure than nonautonomous or dependent learners. However, these considerations of the ideal autonomous learner should be reserved for more mature learners and adult learners, such as undergraduate and graduate students. Quong et al. (2018) found that adult learners, specifically graduate students, were generally more autonomous because they were returning to school by choice and not because they believed they needed a degree to be successful. Additionally, Bergdahl et al. (2020) and Heinrich et al. (2019) found that low and average-performing high school students were more likely to become distracted when using digital technologies by browsing social media or playing games resulting in lower grades while high-performing high school students used digital technologies to support learning resulting in higher grades.

Traditional perceptions of online learners are generally associated with adult learners and are inappropriate for most high school students (Wheatley, 2016). Bollinger and Halupa (2018) wrote "students older than 25 appeared to be more focused on learning outcomes and experiences higher levels of learning than those under 25…the older the student, the more likely engagement was centered on course content and learning" (p. 302-303). When considering younger or less advanced students, such as first-year high school students, there was a consensus amongst researchers that these students were not prepared to be autonomous and successful in online courses where there were diminished interactions with peers and instructors (e.g., Bergdahl et al., 2020; Heinrich et al., 2019; Kontkanen et al., 2017; Larkin & Jamieson-Proctor, 2015; Vickrey et al., 2018). Younger students were also more likely to be disengaged from the course content or distracted by other online tools or websites (e.g., Bergdahl et al., 2020; Heinrich et al., 2019; Heissel, 2016).

It was believed that, in theory, online courses had the potential to provide a better learning experience than a f2f course because of the perceived flexibility, prompt responses and feedback, and media-rich resources (Heppen et al., 2017). "Taking an asynchronous class…requires high levels of self-motivation, self-regulation, and organization" (Protopsaltis & Baum, 2019, p. 32). According to Limtrairut and Marshall (2020),

> In order to succeed in the asynchronous classes, high school students need to be "motivated", "self-directed", "able to work independently", and demonstrate a "willingness to communicate" and to "ask for help". They also need to be "organized" and have "good time management skills". However, at this level they often lack these skills (pg. 8-9).

Motivation and lack of ability to constructively communicate are two of the greatest skills lacking at the high school level, both of which are factors that affect a student's probability of dropping out of an online course (Larkin & Jamieson-Proctor, 2015; Wheatley, 2016; Yates et al., 2021).

Fortunately for the Algebra I student, while learner autonomy may play a part, dialogue and structure have been found to be more significant in lowering transactional distance (e.g., Elyakim et al., 2019; Forte et al., 2016; Yilmaz, 2017; Yilmaz & Keser, 2017). Instructors of Algebra I online courses can help their nonautonomous or dependent learners by making them feel cared for through prompt, personalized responses (Velasquez et al., 2013), by creating activities that allow autonomous and nonautonomous students to collaborate (Elyakim et al., 2019; Yates et al., 2021), or deliberately including strategies designed to help students learn time management and organizational skills to better cope with the online environment (Forte et al., 2016; Murphy & Rodriguez-Manzanares, 2008; Velasquez et al., 2013; Yates et al., 2021).

Key Variables, Designs, and Concepts

Online learning at the secondary level has been researched in great depth in one aspect or another. Most of the research was descriptive and encompassed elements that pertained to the main uses for online learning, student and teacher perceptions and satisfaction in online learning, academic achievement, blended learning, student behaviors, implementation or improving online learning, and communication techniques (e.g., Alfrisa et al., 2020; Belland et al., 2019; Clements, Stafford, et al., 2015; Clements, Zweig, et al., 2015; Hart et al., 2019; Heinrich et al., 2019; Hughes et al., 2015; Kontorovich, 2018; Libasin et al., 2021; Muhonen et al., 2018; Repetto, 2018; Viano, 2018; Yates et al., 2021). Very few studies were focused on comparing online Algebra I student outcomes to f2f, many were focused on credit recovery courses, (e.g., Clements, Stafford, et al., 2015; Hart et al., 2019; Heissel, 2016; Heppen et al., 2017; Hughes et al., 2015), and none were focused on comparing district-run online programs to much larger state-run online programs. Of the few conducted in the last 7 years on f2f and online Algebra I, those by Ahn and McEachin (2017), Heissel (2016), Hughes et al., (2015), and Wakil et al., (2019) were the most similar to this study either in topic or design. The studies by Hart et al., (2019) and Heppen et al., (2017), also compared online to f2f, but for varying reasons, such as credit recovery or likelihood of passing future classes. Of these studies, only Heppen et al.'s (2017) study satisfied all requirements of the WWC's standards for educational research (WWC, 2017b). (This study does not meet the WWC's standards.) Others, such as those by Alfrisa et al. (2020), Choi et al. (2017), Clements, Stafford, et al. (2015), Clements, Zweig, et al. (2015), Heinrich et al., (2019), Kontorovich (2018), Frazelle (2016), Oliver and Kellogg (2015), Rosenzweig et al. (2019), Stevens et al. (2016), and Yates et al. (2021) represented studies that were still concentrated on online high school mathematics, but with varying methodologies, attentions, and purposes. Since failing Algebra I was considered a crucial predictor of whether students go on to graduate high school (Heppen et al., 2017; McGee et al., 2018) the following discussion highlights a few of the studies regarding credit recovery to offer a comprehensive background on what we know so far regarding these at-risk students. Additionally, as noted in Chapter 1, the requested data for state assessments were changed from Milestones assessments to FastBridge aMath assessments and ROI. A brief discussion regarding other studies that had used FastBridge as a measure of achievement is included, followed by synopses of both the closely related studies and other online math research as a background into high school Algebra I courses taken online and to acknowledge where this study has value.

Comparing Online and F2f Algebra I Courses– Studies with Similar Designs, Variables, or Research Questions

Perhaps the most equivalent credit recovery research to this study was a series of studies led by the American Institutes for Research (AIR) and the Consortium on Chicago School Research (Heppen et al., 2017; Rickles et al., 2018). These studies were conducted in Chicago public high schools and included several online and f2f credit recovery Algebra I summer courses. Throughout the series of studies, researchers looked to compare online and f2f credit recovery courses for student outcomes, including class experiences, course grades and test scores, students' math mindsets, students' ability to recover their Algebra I credits, and the likelihood of students earning more math credits in future classes, to graduate, and to get back on track to graduation (Heppen et al., 2017; Rickles et al., 2018). Most of the data collection for these studies were collected over 2 summers and analyzed as necessary to make certain conclusions based on several research questions.

According to Heppen et al. (2017), the experiences perceived by students in the classroom, whether online or f2f, could affect the likelihood of successfully recovering credits. Results indicated that there were no significant differences in student perceived engagement, teacher support, or teacher expectations, however, there was a significant difference in the perceived difficulty and clarity in the course requirements suggesting the online courses were more difficult than f2f (Heppen et al., 2017). In the domain of math skills and mindsets, findings showed significant differences in end-of-course test scores with online students scoring significantly lower. Additionally, how much students "liked" math and had confidence in math for online students indicated significantly less fondness for the subject and less confidence in their math skills (Heppen et al., 2017). Finally, the results for recovering and accumulating credits for graduation indicated that f2f students had a significantly higher likelihood of recovering their Algebra I credit, but that there was no significant difference in whether students would earn math credits in their next math class (typically geometry) (Heppen et al., 2017; Rickles et al., 2018). There was a marginal difference, though not significant, in the likelihood of these students earning credits towards graduation, favoring f2f students (Heppen et al., 2017; Rickles et al., 2018). However, after 2 years, more online students were on track toward graduation, though not significantly (Heppen et al., 2017; Rickles et al., 2018).

The AIR research series was some of the first studies to suggest that online courses were not the recommended route to take for credit recovery students (Heppen et al., 2017; Rickles et al., 2018). It was also one of the very few studies that performed a direct comparison between student outcomes for credit recovery in online and f2f environments making it an essential piece of literature for the field of online education. However, as with most studies, it was not perfect. One of the greatest weaknesses of Heppen et al.'s (2017) study was that the results were based on test scores of students who previously failed a course and were asked to repeat an entire semester, and most likely 2 semesters, worth of material over a 3 to 4-week-long summer course. As noted in

Quong et al. (2018), in that amount of time students may just be getting used to the online environment and learning how to navigate the system without giving as much attention to the assignments as were needed, whereas the f2f students did not have to learn a new technology along with the material and could, therefore, focus on the mathematical content. If the performances of the online students were compared to f2f students' performance after an entire semester or even a year-long credit recovery course, the results might have been more similar.

In Hart et al.'s (2019) study, fixed effects models were used to compare online and f2f "concurrent course performance, future course performance, and likelihood of persisting in high school through the final term of the 12th grade" (p. 2) for both first-time course takers and credit recovery students. Hart et al.'s (2019) findings seemed to imply that students who took virtual courses, whether for the first time or as a retake, had better short-term results as they were more likely to earn better grades than in the f2f classes, for that specific class, but worse mid-term results as virtual students were less likely to take and pass the next course for that subject. Overall long-term results for math students seemed to differ for the two groups as first-time course takers were less likely to graduate while credit recovery students were more likely to graduate.

Hart et al.'s (2019) study closely replicated that of Heppen et al.'s (2017) study but included a broader range of topics considered for the two environments and a different demographic, resulting in contrasting findings. Heppen et al.'s (2017) population were mostly students from lower-income families in Chicago where 86% of the participating students qualified for free or reduced lunch, and the top three ethnicities were 57% Hispanic, 33% African American, and 8% White. In Hart et al.'s (2019) population approximately 40% of students qualified for free or reduced lunches, and the top three ethnicities were approximately 50% White, 22% Black, and 24% Hispanic. Research has shown that, regarding Black versus White students and wealthy versus impoverished students, White students tended to outperform Black students in mathematics and wealthy students tended to outperform impoverished students (e.g., Easton, et al., 2017; Paschall et al., 2018). Based on demographics alone, it was not surprising that Hart et al.'s research suggested students were more likely to perform better online than in a f2f course as Hart et al.'s students were mostly wealthier, White students. Had both studies been conducted with more similar populations, the results of Heppen et al.'s (2017) and Hart et al.'s (2019) studies may have been more alike.

Similar to Heppen et al.'s (2017) and Hart et al.'s (2019) studies, Heissel's (2016) research examined academic achievement for students taking Algebra I virtually and f2f but utilized the outcomes for eighth-grade Algebra I virtual course takers and compares them to the outcomes of both 8^{th} and ninth-grade f2f course takers. Comparisons were made for both a specific district and separately for the state. Following a North Carolina policy change in 2011, Algebra I became a class that advanced eighth-grade students could opt in and take virtually. Those who did not opt to take it virtually could still take the course f2f, in most counties. Also, those who were not advanced stayed on the traditional track and did not take Algebra I until ninth-grade. While the virtual Algebra I class was optional and although these students were advanced eighth-grade students, Heissel's (2016) results indicated that, in both the specified district and overall in the state, these students did not perform as well in Algebra I as the f2f eighth or ninth-grade students overall, and that the eighth-grade virtual students had Algebra I end-of-course test scores that were 6 to 10.3 percentage points lower than students who took the course f2f, though this did not mean these students failed. Academic achievement, including end-of-course test scores, for students that waited until ninth grade tended to be higher than their eighth-grade peers. Heissel (2016) suggested that the lower-than-expected results for the advanced eighth graders and higher-than-expected results for lower eighth graders could be due to the virtual program having "course content [that is] specific and narrow, improving outcomes for relatively lower-performing students but limiting the knowledge gained by more able students" (p. 112), or from a "spillover effect" in which the lower performing eighth graders received more one-on-one time with teachers after the advanced classmates were no longer in the same class.

Further analyses of Heissel's (2016) research showed that roughly 58% of eighth - grade students took Algebra I virtually while about 25% took the course f2f (the remaining students took the course in ninth grade). Results regarding the virtual courses showed a 33% dropout rate. Heissel (2016) explained that "48 students officially dropped the NCVPS course, and 310 students enrolled in NCVPS but never took the test, implying a dropout rate of 33%" (p. 101). This equates to approximately 15% of students officially dropping the course. Heissel (2016) also suggested that 14% of the enrolled students took the end-of-course test in 2012, implying that these students were accidentally enrolled in virtual Algebra I, but never actually in the course. The last 14% of dropouts were estimated to also be errors or students who moved, however, "many teachers noted that non-test-takers often failed to complete assignments throughout the class" (Heissel, 2016, p. 102). While Heissel (2016) considered these students as not "true members of the course" (p. 102), other research has suggested that these students be considered dropouts, withdrawals, or incompletes if they participated in the course for some pre-determined amount of time (Akos & James, 2020; Hassan et al., 2019).

While dropout rates were not calculated, Wakil et al. (2019) and Ahn and McEachin (2017) found similar results to Heppen et al., (2017) and Heissel (2016), that virtual students did not perform as well as f2f students. Wakil et al.'s, (2019) research participants included seventh and 10^{th}-grade students in Iraq taking statistics either in a f2f classroom or virtually via Skype. Results were based on end-of-course test scores (each course only lasted 6 weeks) and revealed that the f2f pass rates were 91% and 83% for seventh-grade students and 81% for 10^{th}-grade students. The virtual course counterparts had a pass rate of 67% for seventh grade and 60% for 10^{th} grade. This study's population, in ethnicity and grade level, was quite different from the others mentioned so care must be taken when making comparisons between the studies, however, the methodology of using end-of-course test scores to make comparisons between f2f and virtual classes was consistent with each study.

Ahn and McEachin (2017) examined student records for traditional public schools, traditional charter schools, and full-time online schools to assess patterns in subgroups of students and the academic achievements of the students in each environment. Student data included in the study covered students from kindergarten through 12^{th} grade for the school years 2009-10 to 2012-13. Student subgroups included race, free and reduced lunch, gifted education, special education, English as a second language, attendance and suspensions, grade levels, and home districts. Ahn and McEachin (2017) found that e-school enrollments increased by approximately 60% over 4 years while traditional charter school enrollments increased by about 6% and traditional public-school enrollments decreased by approximately 5% over the same time frame. Results for the demographics of each environment indicated that public school students at the high school level had higher prior academic achievement on average than both the online schools and charter schools at the same level, as well as more students in gifted education, fewer disciplinary actions, fewer students in special education, fewer students repeating a grade, and fewer students that qualified for free or reduced lunches (Ahn & McEachin, 2017). Charter schools had a higher percentage of Black students (60%) than online schools (10%) and public schools (12%); more English language learners, more Hispanic students, more students with lower prior academic achievement (on average), and fewer White students than the other two environments. Online schools carried the highest enrollments for White students and the lowest enrollments for males, suggesting there were higher numbers of White females enrolling in online schools.

When evaluating student academic achievements, Ahn and McEachin's (2017) calculations showed that students in online and charter schools were less likely to pass the Ohio Graduation Test (OGT) on their first attempt for all subjects (math, reading, sciences, social studies, and writing) than students in traditional public schools. Students in online schools had lower achievement and were significantly less likely to pass their OGTs than students in charter schools, especially when students' prior achievement levels put them in the lower two-thirds of overall achievement distributions (Ahn & McEachin, 2017). These results suggested that students in online schools may have been making far fewer academic gains than students in public or charter schools, putting them behind, academically, than their peers in other environments, even if they were in the upper one-third of overall achievement distributions before attending school online.

In Hughes et al.'s (2015) research, student outcomes were compared for online and f2f courses for general education and credit recovery courses but then were also analyzed based on student demographics. Closely related to this study, Hughes et al.'s (2015) study used logistic regression analysis and student test scores to examine the likelihood a student had of earning a grade of C or higher in the respective course environments.

The data for Hughes et al.'s (2015) study was collected from the Florida Department of Education's Education Data Warehouse through de-identified student transcripts for the 2007-08 through 2010-11 school years. Hughes et al.'s (2015) study included analyses for both the 20 most common general education courses and the 20 most common credit recovery courses, however, only the ninth-grade information was discussed here as it most closely relates to this study.

According to Hughes et al. (2015), students taking online classes were more likely to be White and less likely to be in special education, English language learners, or eligible for free or reduced lunch. For the top 20 most common online courses, students who scored 3 or higher on a previous Florida Comprehensive Assessment Test (FCAT) were 7.3 percentage points (pp) more likely to earn a C or better in online classes than students taking the same course f2f for ninth-grade students (80.4% percent likelihood in f2f) down to -0.1pp for 12th-grade students (89.5% likelihood in f2f) indicating that f2f students were slightly more likely to earn a C or better. For students who scored lower than 3 on a previous FCAT, the likelihood of performing better in an online course rose to 14.9 pp for ninth-grade students (64.5% likelihood in f2f), then decreased to 3.1pp for 12th-grade students (79.1% likelihood in f2f). The likelihood of earning a C or better seemed to resonate throughout the various demographic analyses as well with likelihoods being higher for online learners than f2f. The demographic subgroups considered by Hughes et al. (2015) included White, Black, Hispanic, eligible for the school lunch program, in special education, and English language learners. Additionally, Hughes et al.'s (2015) study indicated that "the success gap between online and f2f courses was 36.6 percentage points in grade 9, the largest difference found in the study" (p. 12).

Hughes et al. (2015) specified that this study did not indicate that students in online courses did or will do better than those in f2f courses. First, students were given the option to take the course online or f2f, and those who took the online version were possibly students who felt more comfortable with technology and had better self-management and self-motivational skills. Second, the students who chose to take online courses had, in general, scored higher on the previous year's FCAT than the students who chose to take the course f2f, and "students with higher FCAT scores are probably also more likely to earn higher grades" (Hughes et al., 2015, p. 7). Third, Hughes et al. (2015) grouped the results for all courses, either online or f2f, as one unit instead of analyzing the odds for each course. It was possible that, for example, the odds of earning a C or better in f2f Algebra I were much higher than online, but the results of all the other courses in addition to Algebra I skewed the results to be in favor of the online versions. Fourth, the results of Hughes et al.'s (2015) study were based solely on the end-of-course grades, which the authors admit were highly subjective. For example, the quality of a student's writing, and therefore their grade, was highly dependent upon a teacher's preference for and methods for scoring essays. A more nonsubjective method for comparing student outcomes would have been to consider end-of-course exams or state assessments.

Ahn and McEachin (2017), Heppen et al. (2017), Hart et al. (2019), Heissel (2016), Hughes et al. (2015), and Wakil et al.'s, (2019) studies were a few of the reports conducted in the last 7 years that had a concentration on comparing online and f2f student outcomes for secondary math students. While Heppen et al.'s (2017) study focused solely on Algebra I credit recovery courses, Hart et al. (2019) and Hughes et al.'s (2015) research grouped semester or year-long courses into bulk analyses instead of focusing on individual courses, such as Algebra I specifically. Even with the results supplied by Ahn and McEachin (2017), Heissel (2016), and Wakil et al., (2019), there was still a great need to consider the student outcomes for semester or year-long Algebra I courses taken online and f2f.

FastBridge Assessments as a Tool for Understanding Achievement

The present study requested results of the end-of-course state assessments known as Milestones. However, a recent program change meant results were received for FastBridge's aMath assessments instead. A Google Scholar search for "FastBridge aMath assessment" returned 85 results with a little over half mentioning FastBridge directly (others were returned due to phrases such as "a math assessment"), and almost all that did were recent Master's Capstones or Ph.D. Dissertations. Only three papers were found that were written by nondegree-seeking authors. In one paper, one of the authors was noted as having "equity and royalty interests in, and served on the Board of Directors for, FastBridge Learning (FBL) LLC" (Albano et al., 2018, p. 179), but this study was intended to produce "a new method for evaluating equating in longitudinal contexts" and had no mention of using FastBridge assessments. The other two papers by nondegree-seeking authors were written by individuals within the Illuminate Education organization and are discussed below. Of the available studies, those that focused on reading, social-emotional learning, behavior, development, CBMmath, or earlyMath, all of which are available assessment and screening programs from FastBridge, were excluded from review as they were not relevant to the current study. Additionally, studies whose focus was on younger elementary students (kindergarten to sixth grade) were not reviewed unless directly tied to FastBridge aMath scores. Even as limited as the reports were, a look into how these other studies used FastBridge data could prove useful in understanding the benefits and downfalls of basing achievements on these assessments.

FastBridge's computer-based Universal Screening tool is a combination of FastBridge's aMath, Reading, CBMmath Automacity, CBMmath Concepts and Applications, and mySAEBRS, and is used as a screening tool to identify students who may need extra support in reading, math, and behavioral accommodations (Illuminate Education, 2022a). In Carr's (2021) study, the Universal Screening tool was used to determine if there was a difference in achievement from the winter to the spring dependent on whether a student attended an after-school program. Students were African American, White, and Hispanic males in sixth to eighth grade. Carr (2021) used the winter and spring raw scores for reading and math along with a mixed-model ANOVA to analyze the main and interaction effects of the after-school program and race on students' scores. Significant differences in math scores from winter to spring were found, as were significant differences in scores between the races. However, no differences were found between African-American students who attended the after-school program, and those who didn't.

Biery (2022) used aReading, aMath, and mySAEBRS scores for the fall and winter semesters to understand the relationships between an online mindfulness intervention and academic achievement, attendance, behavior, and social-emotional growth for ninth-graders. A one-way ANOVA was used to compare aMath scores from the year prior to the mindfulness intervention to the first year the mindfulness intervention was used. Biery (2022) used raw scores for Fall 2020, Winter 2020, Fall 2021, and Winter 2021 assessments, and ROI for the Fall to Winter of 2020 and Fall to Winter of 2021. Results indicated that in the year in which the mindfulness intervention was not used, 2020, students had higher averages on both the Fall and Winter assessments and had overall more growth between assessments, though neither were statistically significantly different.

In a similar study by Straub (2021), aMath and aReading scores were used to examine if there was a correlation between the FastBridge assessments and the Kansas

Assessment Program's (KAP) English Language Arts and math scores for students in Grades 3-8. Straub (2021) used raw scores from the Winter and Spring FastBridge and KAP assessments for the 2018-2019 school year. Results indicated strong, positive, statistically significant relationships for both math and ELA for each testing period. Additionally, Straub (2021) found that the spring aMath and KAP Math assessments had a stronger correlation than the winter assessments.

The last two studies regarding FastBridge assessments were funded by Illuminate Education and were published 5 months apart in August 2020 and February 2021. These two reports look at the learning loss and academic achievement gaps following COVID-19 school shutdowns. In the 2020 study, Bielinski et al. used scores from aMath assessments to determine average learning loss from spring to fall in a normal year and estimated the average learning loss to occur in the fall of 2020 after students had been out of school for more than just the summer break. Results indicated that "math achievement loss as measured by aMath was observed across all grades from Kindergarten through Grade 5" (Bielinski et al., 2020, p. 12). Bielinski et al., (2020) suggested that schools be prepared for students to be well behind where they would typically be in fall with reading skills being 1 to 2 months behind and math skills being 2.5 to 4.5 months behind.

In Bielinski et al.'s (2021) second study, FastBridge aMath scores and ROI were used to understand achievement growth, or decline, between the fall of 2019 and the fall of 2020. The authors noted "using growth instead of achievement controls for random variation due to changes in demographic composition from year to year, and using the ROI controls for the variation in the interval between administrations" (2021, p. 5).

Results indicated that learning growth during the year of the COVID-19 shutdowns was lower than an average year's growth and that the difference between the average year gain and the COVID year gain exceeded the standard error, especially in math. While effect sizes were found to be relatively small, Bielinski et al. (2021) estimated about 3 to 4 months of loss from Grade six to Grade seven and approximately a 3 month loss from Grade seven to Grade eight. The authors predicted that a 3 month loss in the upper grades would need a sustained gain of about 33% over a typical school year to make up for the learning loss during the pandemic. Additionally, the authors compared the results of their study using FastBridge with the results of other studies that used Measures of Academic Progress (MAP), i-Ready, and STAR. Slight differences were found between aMath and MAP as well as aMath and i-Ready, but similar results were found between aMath and STAR. Further results "indicate that the drop in performance during the COVID year is nearly uniform across school types, race/ethnicity…and poverty levels" (Bielinski et al., 2021, p. 13). Recommendations for getting students back on track included continued screening using FastBridge assessments, using ROI in addition to benchmark goals for getting students back on track, spending more time on math than other subject areas, and ensuring remote instruction is effective.

In each of the mentioned FastBridge studies, results from aMath assessments were used to measure academic achievement or growth (ROI). Carr (2021) looked at achievement and an after-school program, and Biery (2022) considered achievement due to an online mindfulness intervention much like the present study considered achievement due to the learning environment (online and f2f). Bielinski et al. (2020,

2021) used FastBridge aMath scores to predict learning outcomes and growth due to COVID-19 school shutdowns but also compared aMath results with those of other commercially developed screening programs much like the work of Straub (2021). While the present study did not attempt to compare FastBridge aMath results with other assessment programs, it was welcoming to note that aMath was a valid and reliable tool for assessing student achievement and learning growth. Further details as to the specifics of FastBridge aMath assessments are provided in Chapters 3 and 4.

Comparing Online and F2f Algebra I Courses – Studies with Varying Designs and Approaches

In parallel studies, Stevens et al. (2016) and Frazelle (2016) investigated online credit recovery courses taken in Montana during the 2013/14 school year. Stevens et al. (2016) investigated the enrollments and performance trends throughout all the Montana high schools that used the Montana Digital Academy (MTDA) for their online credit recovery courses. According to Stevens et al. (2016) report, during the 2013/14 school year, MTDA had 2,452 students enrolled in more than 50 different credit recovery courses totaling 3,763 enrollments (some students enrolled in more than one course). Most of these enrollments were for English language arts (37%), then nearly equally divided between math (19%), social studies (19%), and science (17%) (Stevens et al., 2016). Within these enrollments, there were more male (60%) than female students, and most students were in either Grade 10 (34%) or Grade 11 (31%) (Stevens et al., 2016).

Stevens et al.'s (2016) research also closely reviewed the passing rates of different subgroups within the MTDA enrollments for this school year. The overall passing rate for the 2013/14 school year was 57% (Stevens et al., 2016). According to Stevens et al. (2016), of the 43% who did not pass their course, “all but five dropped out of their course before receiving a grade” (p. 5, 7). Of those who passed their courses, there were more females (60%) than males, and most were in Grade 12 (63%) (Stevens et al., 2016). Stevens et al. (2016) also found that students who took more than one course during a semester were more likely to pass than students only taking one course. Passing rates amongst students taking one, two, three, or four courses were 40%, 68%, 82%, and 85%, respectively (Stevens et al., 2016). Additionally, passing rates were the lowest for math courses at just 49% while the highest passing rates were in social studies at 71% (Stevens et al., 2016).

Frazelle’s (2016) research purpose was to gather strategies from the top-performing Montana high schools on how they created successful online credit recovery courses, Frazelle’s (2016) study was conducted through interviews with six online credit recovery facilitators from different high schools in Montana. Five of the six schools had online student pass rates that were higher than the state median of 60% for the 2013/14 school year (Frazelle, 2016). The sixth school in the study had an online student pass rate of 51% but also had enrollment numbers higher than most other high schools (Frazelle, 2016). Frazelle’s (2016) interviews asked facilitators to comment on their perspectives as to the strategies they believed contributed to their school’s high pass rates. Responses were organized into one of four clusters and then further assorted into more specific categories. Frazelle’s (2016) results suggested that the four most indicated strategies included:

1. Establishing a consistent program structure by ensuring students have a specific location and time during regular school hours to complete online coursework, making sure students have enough time during the school day to complete work, providing a certified teacher to serve as the online program facilitator, and securing external funding sources if needed.
2. Providing instructional support by aiding students in creating course goals and following up with students on a frequent and regular basis, identifying other teachers within the building who students can go for help if needed, encouraging students to communicate with their online instructors, and promoting staff responsibility for student success while removing barriers and opportunities for student excuses.
3. Building rapport between students, facilitators, and online instructors.
4. Utilizing tools and resources to monitor student progress beyond student grades online. Items to track should include basic student information as well as adult contact information and online activity.

According to Clements, Stafford, et al. (2015) and Clements, Zweig, et al. (2015), one of the most popular reasons for offering online education was to provide students an alternative opportunity to recover credits for courses failed in the f2f classroom. Clements, Stafford, et al. (2015) and Clements, Zweig, et al.'s (2015) studies included a series of surveys regarding their online learning programs sent to a sampling of New York high schools and Iowa and Wisconsin school districts. In the New York high schools, 82 percent of those that responded and said they offered online courses claimed that providing credit recovery courses online was an important reason for doing so (Clements, Zweig, et al., 2015). Similarly, credit recovery was reported as being very important and one of the main reasons for offering online learning by responding school districts in Iowa (71%) and Wisconsin (66%) (Clements, Stafford, et al., 2015). Math courses were reported as one of the most commonly taken courses online with Iowa reporting 73% of schools offering math courses online, and Wisconsin reporting 81% of schools (Clements, Stafford, et al., 2015).

The Clements, Stafford, et al. (2015) and Clements, Zweig, et al. (2015) surveys also included items regarding administrator concerns and challenges when implanting an online program. In all three areas surveyed, course quality was one of the top concerns reported in New York (reported by 71% of schools) and Wisconsin (reported by 37% of school districts) and came in as the second most cited concern in Iowa (reported by 32% of school districts) (Clements, Stafford, et al., 2015; Clements, Zweig, et al., 2015). For New York schools, academic integrity and lack of f2f interaction were the next two most cited concerns (Clements, Zweig, et al., 2015), while Wisconsin school districts' results indicated a lack of funding and lack of student interest (Clements, Stafford, et al., 2015). The top concern for Iowa schools was the lack of teacher training, reported by 61% of the school districts (Clements, Stafford, et al., 2015).

In the Iowa/Wisconsin study, additional survey items included questions regarding how students' progress was tracked and the role of onsite monitors. Results indicated that just over half of the school districts that had online programs always assigned an onsite monitor while Wisconsin reported always having an onsite monitor in over 75% of their online programs (Clements, Stafford, et al., 2015). When it came to monitoring student progress, both Iowa and Wisconsin reported the students' final grades were the main means of monitoring progress with 85% and 90%, respectively (Clements, Stafford, et al., 2015). Completion of assignments was the third most reported means of monitoring student progress for both sites while tracking attendance and time spent online were the least two used methods for both Iowa and Wisconsin (Clements, Stafford, et al., 2015).

Research surrounding online education is becoming more abundant each year and varies greatly in focus. While Clements, Stafford, et al. (2015) and Clements, Zweig, et al. (2015) investigated the types of courses being taken, and Frazelle (2016) inquired into strategies that make online learning successful, Oliver and Kellogg (2015) explored possible reasons for student failures leading to the need for credit recovery. Oliver and Kellogg's (2015) study was a mixed methods design comprised of a Likert scale and open-ended survey questions presented to both teachers and students of a state-sponsored virtual school. Virtual courses offered included credit recovery to advanced placement courses in math, English, social studies, and science. 193 credit recovery students (of 862 total student-submitted surveys) and 36 credit recovery teachers (of 128 total teacher-submitted surveys) responded to questions pertaining to reasons for taking an online course, typical grades received in school, why students failed the original course, why the credit recovery courses were better, and course design. Key findings indicated that just over 50% of students said they failed originally due to self-discipline issues (37.3%) or poor teaching (14.5%); 77 students (about 40%) said that online courses were better because they could self-pace, the teachers were more involved, and they had more time to complete assignments; and, at the time of the surveys, only 61.4% (divergent standard deviation of 19.6%) of teachers believed their credit recovery students would pass their online classes. Two of the greatest concerns with this study, however, was that the virtual school staff were highly involved in the writing and editing of the survey questions which may have led to bias in the wording of the questions, and that there was no attempt to compare final grades received for online students, so success was calculated based on student and teacher perceived success.

Beyond the research on online credit recovery courses, other research regarding online mathematics can provide further background information to establish the necessity of the current research. The rest of this section will review research by Alfrisa et al., (2020), Choi et al. (2017), Heinrich et al., (2019), Kontorovich (2018), Rosenzweig et al. (2019), and Yates et al. (2021).

Choi et al. (2017) considered the effects of self-reflection on overall math performance for elementary to high school students in online learning environments, and whether students engaging in self-reflection affected online math performance. A pilot study was conducted prior to an extended study. Their data came from a set of assessments given to students at the end of each lesson and/or unit. For the pilot study, students were given two questions that asked them to rate their understanding of the material and confidence with the topic on a four-option scale. The extended study included four questions, still on a four-option scale, regarding “(a) general feelings towards math, (b) the use and preference of learning strategies, (c) self-judgement of skill level, and (d) identifying skills as strengths and/or weaknesses" (Choi et al., 2017, p. 84). Math performance levels were based on unit tests and overall course grades.

Results of the self-reflection assessments during the pilot study showed that students accurately judged their understanding of the material and their confidence levels compared to their actual math performance and that confidence levels were higher for students who pretested higher or attended the same school the previous year (Choi et al., 2017). Results for the extended study indicated that 70% of students agreed or strongly agreed with having positive feelings about math, students preferred to use certain learning strategies, and that final course scores were positively correlated with the number of times students took reflection assessments. The elementary and middle school students were found to have taken more reflection assessments than high school students, and generally answered all of the questions. Middle school students tended to have significantly higher unit test scores when they completed unit reflection assessments. However, in high school, a high percentage of students took every reflection assessment early in the year, but as the year progressed and the material became more difficult more than half of the students stopped completing the assessments. For high school students, students who completed reflection assessments on more difficult units had lower unit test scores than students who did not complete the assessments (Choi et al., 2017). Overall, Choi et al.'s (2017) study indicated that elementary and middle school students were more likely to engage in self-reflection activities and that self-reflection was linked with improved online math performance, but high school students were less likely to engage in self-reflection activities and tended to perform lower when they did, especially on more difficult topics.

A similar study to Choi et al. (2017) was that of Rosenzweig et al. (2019). Instead of self-reflection assessments as an intervention, Rosenzweig et al., (2019) considered whether students' utility value or perceptions of course usefulness affect their online math course performance, specifically for Algebra I and Geometry. In this case study, three types of utility value interventions were created and implemented. The first intervention asked students to write an essay about how math related to their lives. The second intervention asked students to read and evaluate quotations about the value of mathematics. The third intervention combined the first two and asked students to first read quotations about the usefulness of mathematics and then respond with an essay. While attrition rates for the three interventions ranged from 24.7% for the essay-only intervention to 36.8% for the quotes and evaluation intervention, Rosenzweig et al. (2019) found that students who engaged in utility value interventions that required them to read quotations and evaluate them had the highest posttest utility values. The two interventions that required essays had no significant differences in posttest utility values. A major pitfall of this research, recognized by the authors, was that only 7% of the students who were invited to participate did, and "it is possible that more engaged, self-efficacious, and interested students chose to [participate]… this self-selected sample might be more likely to benefit from interventions than other students because they would have been more likely to engage deeply with intervention materials" (Rosenzweig et al., 2019, p. 347).

In similar studies, Kontorovich (2018) and Alfrisa et al., (2020) considered the effects of utilizing online tools in a f2f high school Algebra course (also called blended learning) on student performance and achievement. Alfrisa et al.’s (2020) report was a systemic review of previous blended learning research conducted from 2014 to 2018. Based on the results of ten other studies, Alfrisa et al. (2020) concluded that blended learning was very effective for the learning of high school Algebra, especially when it used the strengths of f2f and online learning to overcome obstacles often seen in Algebra classes.

Likewise, Kontorovich (2018) investigated how the use of online asynchronous discussion forums correlated with student achievements in f2f high school Algebra courses. Kontorovich’s (2018) research found that students who consistently initiated threads outperformed students, on quizzes and the final exam, who only replied to the threads of others, and that students who were active in their threads were more likely to be active in the threads of other students. Additionally, Kontorovich (2018) sorted the types of threads being initiated into categories in which students asked for help getting a problem started, understanding what a question was asking, or clarification on class materials; asked for help or verification on solutions they hadn’t yet started, started but got stuck, or finished; or to discuss inferences about the material that were not discussed in class.

Kontorovich’s (2018) findings were consistent with Alfrisa et al.’s (2020) review and that of TDT. In Kontorovich’s (2018) study, student-student dialogue was high rather than student-teacher, yet transactional distance appeared to be lowered, especially in the case of the most active students, due to a high number of interactions.

Heinrich et al.'s (2019) 4-year mixed methods study used a series of online student observations, interviews with administrators and support staff, and student records matched with the online vendor's records to quantify student demographics and online course-taking patterns. Results on student demographics for the district suggested that students taking online courses were mostly Black students, from low-socioeconomic backgrounds, had failed a course the previous year, pregnant or parenting students, students returning from incarceration or expulsion, students with lower average math and reading test scores than the district average, and students with more absences than the district average. Failing a course drastically increased a student's odds of taking the course again the following year, and "students who failed a course in the prior year had 126% higher odds of taking a course online, affirming the "credit recovery" focus of online instruction in the district" (Heinrich et al., 2019, p. 2165).

Heinrich et al., (2019) grouped students taking online courses into four categories of users (engaged users, moonlighters, nominal exerters, and incompatible users) depending on their level of engagement with the course either in school or at home. Engaged users were the most engaged and each group after that was close to 50% less engaged than the one before it, except for moonlighters. 11^{th}-grade students made up the majority of the engaged students and moonlighters. Moonlighters were often as engaged in their courses as the engaged learners but spent 80% or more of their time online outside of school. "The incompatible users took only one course on average, spent the least amount of time in their course and online sessions, and completed the fewest number of activities in their courses per day" (Heinrich et al., 2019, p. 2168). 9th and 10th graders made up more than 60% of the incompatible users and over half of them were taking math or reading. Unsurprisingly, Heinrich et al., (2019) found that students who were engaged with their online courses were more successful in their online courses than students who were incompatible and less engaged. Each additional percentage point of idle time users spent in their online courses accounted for one-third of a percent less likely to pass the course, one-fifth of a percent less likely to pass on time, and course grades being about 0.42 points lower. Additionally, Heinrich et al., (2019) found that students who took online courses in high school were increasingly more likely to have lower average GPAs and test scores, especially in math and reading, for each additional year of online courses taken versus students who did not take online courses.

From interviews with teachers and administrators, Heinrich et al. (2019) reported the greatest concerns with online course taking for high school students. Among the most reported concerns were low reading levels of online students, lack of ability to accommodate students with special needs, lack of on-site teacher or aide with content expertise (especially for math), enrollment of students just needing a place to be throughout the day, high student-teacher ratios (45:1 up to 65:1 was reported), lack of resources, lack of student engagement even when encouraged to participate, and lack of active engagement versus passive engagement. Heinrich et al. (2019) reported that during observations many students were seen clicking links or acting engaged with the content, but also were talking with others or playing on their phones while the course content played in the background. Instead of learning the material, students would look up answers to assessment questions online, then copy and paste them into the solution areas, raising the question as to whether online courses were used to truly get students to content mastery or “simply to provide students an opportunity to earn credits needed to graduate” (Heinrich et al., 2019, p. 2178).

The COVID-19 pandemic greatly affected the way we view online learning as students worldwide were forced into the virtual environment due to f2f school shutdowns. Yates et al. (2021) conducted a mixed methods study to examine how high school students in their last 2 years of high school experienced learning from home during COVID-19. Through surveys sent to students aged 16 and older, Yates et al. (2021) were able to decipher students’ perceived amount of learning gained, preferences for online or f2f learning, perceived flexibility of the online courses, perceived collaborative and authentic learning experiences, online learning motivation, and preferred pedagogy in the online environment. Results indicated that among students that spent less time on schoolwork at home than in the classroom, 66% reported learning less at home. Among students that spent more time on schoolwork at home, 36% reported learning less at home while 35% reported learning more. When considering course flexibility and preferred learning locations, 90% of students preferred learning in the classroom, but wanted to see the continued flexibility with pacing and digital resources continued in the f2f environment. Authentic learning experiences that considered students’ COVID-19 lifestyles were reported by some students, however, most students reported that learning experiences while at home were the same as they would be in school. Additionally, students reported that their preferred learning activities included direct instruction, pre-recorded lessons made either by the teacher or found online, and competitive activities such as games. The lack of authentic learning experiences, lessons that did not include students' preferred learning styles, as well as self-reported lack of time management skills, may have been contributing factors to the nearly 40% of students who cited a lack of motivation as one of the hardest parts of online learning.

Yates et al. (2021) found that 53% of students preferred collaboration in the classroom where they could receive immediate feedback from the teacher and classmates. Students who preferred collaborating in the classroom reported that collaborating via videoconferencing apps was difficult because only one person could talk at a time, the teacher would dominate the conversations, or classmates would turn off their cameras and microphones and not participate. 67% of students who claimed online collaboration was not helpful also reported that they learned less at home while 52% of the students who claimed that online collaboration was helpful reported that they learned more at home. These findings are consistent with the dialogue aspect of TDT.

Summary

In recent decades online learning has become ever-more prevalent, as has online credit recovery. In its earliest forms, distance education was nothing more than postal correspondence between a teacher and student (Moore, 1997). During these times, transactional distance was considered quite high, especially due to the lack of dialogue and rigid structure (Moore, 1997). As distance education evolved into online learning, learner autonomy became an element of the TDT (Limtrairut & Marshall, 2020; Moore, 1993, 1997; Saba & Shearer, 2018). TDT has since become a vital part of understanding online learning (Dockter, 2016; Elyakim et al., 2019; Protopsaltis & Baum, 2019; Velasquez et al., 2013; Yilmaz, 2017; Yilmaz & Keser, 2017) and has proven to be helpful when studying high school mathematics in an online environment. While other theories, such as TPACK (Koehler & Mishra, 2005) or Engagement Theory (Kearsley & Schneiderman, 1998) can provide additional insight into understanding online credit recovery programs, recognizing the interplay of how dialogue, structure, and learner autonomy can affect student learning in both online and f2f environments may provide some of the essential viewpoints for making comparisons in the achievements of students in both settings.

TDT plays a vital role in the development of both online and f2f high school mathematics classes. Whether dialogue was f2f or online, asynchronous or synchronous, having more high-quality interactions between student-student or student-teacher can improve engagement, motivation, and perceived learning (e.g., Dumford & Miller, 2018; Protopsaltis & Baum, 2019; Quong et al., 2018; Weiss & Belland, 2018). Frequent and timely feedback, especially in online environments, was necessary for students to better understand the material, navigate online platforms, maintain engagement, and better deal with feelings of isolation (e.g., Frazelle, 2016; Viano, 2018; Wheatley, 2016).

As stated earlier, students were more satisfied in their online courses when learning occurred in interactive environments that allowed for the construction and sharing of knowledge (e.g., Grigoryan, 2017; Yilmaz, 2017). Interactive environments that allowed for collaboration were part of having a structure that lowered transactional distance (Moore, 1997), as was creating a course structure in which students had some choice over what was learned and there was flexibility as to when assignments were due (Yates et al., 2021). Unfortunately, subjects that required more regulated sequencing, such as most mathematics courses, did not allow for as much flexibility (Moore, 1997; Usta & Mirasyedioğlu, 2022).

Having an optimal course structure and opportunities for dialogue does not mean transactional distance will be lowered. Students must have some sense of learner autonomy, such as motivation to engage in and stay focused on the course materials or class discussions, in order to lower transactional distance (e.g., Bergdahl et al., 2020; Heinrich et al., 2019). However, motivation and lack of ability to constructively communicate are two of the greatest skills lacking at the high school level, especially in ninth-grade students (e.g., Heinrich et al., 2019; Oliver & Kellogg, 2015; Wheatley, 2016; Yates et al., 2021). Usta and Mirasyedioğlu (2022) noted that even at the university level, students taking online algebra stated, "the factors that hinder their learning as lack of communication, inability to work together, affective and psychological reasons" (p. 162), and that "the practice of distance education is not effective in increasing students' success in solving algebra problems" (p. 163). Transactional distance was observed to be higher in the online courses than in the f2f courses (Usta & Mirasyedioğlu, 2022). Should the f2f students of this study outperform the online students, it was reasonable to suggest that the online students may have faced difficulties associated with greater transactional distance, impairing their learning outcomes, as was the case in the previously mentioned studies (Bergdahl et al., 2020; Usta & Mirasyedioğlu, 2022).

In addition to TDT research, recent trends in online high school mathematics research have been to evaluate how online learning is being used (e.g., Alfrisa et al., 2020; Clements, Stafford, et al., 2015; Clements, Zweig, et al., 2015; Heinrich et al., 2019; Hughes et al., 2015), student and teacher perceptions and satisfaction in online learning (e.g., Clements, Stafford, et al., 2015; Clements, Zweig, et al., 2015; Frazelle, 2016; Heinrich et al., 2019; Oliver & Kellogg, 2015; Rosenzweig et al., 2019; Yates et al., 2021), academic achievement (e.g., Ahn & McEachin, 2017; Alfrisa et al., 2020; Hart et al., 2019; Heinrich et al., 2019; Heissel, 2016; Heppen et al., 2017; Hughes et al., 2015; Kontorovich, 2018; Wakil et al., 2019), student behaviors (e.g., Heinrich et al., 2019; Kontorovich, 2018), interventions for improving online learning (e.g., Choi et al., 2017; Frazelle, 2016; Kontorovich, 2018; McKenzie et al., 2022; Yates et al., 2021), and communication techniques (e.g., Kontorovich, 2018; Libasin et al., 2021; Muhonen et al., 2018; Viano, 2018; Yates et al., 2021; Yilmaz & Keser, 2017;). Although online courses may be vital to a student's chance of graduating high school, especially for credit recovery students (e.g., Clements, Stafford, et al., 2015; Clements, Zweig, et al., 2015; Frazelle, 2016; Heinrich et al., 2019; Heppen et al., 2017; Stevens et al., 2016; Viano, 2018), most current research has been focused on how particular interventions affect student learning in online courses (e.g., Alfrisa, et al., 2020; Choi et al., 2017; Kontorovich, 2018) or student perceptions of online learning following the COVID-19 era school shutdowns (e.g., Yates et al., 2021).

Research comparing online learning to f2f has had mixed results on whether online education was as effective as the f2f environment in getting students to content mastery to pass end-of-course tests and move closer to graduation. While some studies reported findings of students in online courses performing better than students in f2f courses (e.g., Hart et al., 2019; Hughes et al., 2015), others maintained that online students had lower test scores, lower GPAs, and lower overall academic gains (e.g., Ahn & McEachin, 2017; Heinrich et al., 2019; Heissel, 2016; Heppen et al., 2017; Wakil et al., 2019). According to Hart et al. (2019) and Hughes et al. (2015), the students in the online courses that were performing better than students in the f2f courses were White students and students who did not qualify for free and reduced lunches. Ahn and McEachin (2017) also claimed that the majority of online students were White, however, their study found that these students did not perform as well as students in f2f courses. Heppen et al. (2017) and Heinrich et al. (2019) both suggested that the students in the online courses that were not performing as well as the f2f courses were mostly Black or Hispanic students and students that qualified for free or reduced lunches. Between these five studies, written within 4 years of each other, there were conflicting reports as to who the majority of online students were and whether online students performed better or worse than f2f. The differing outcomes and reported majority groups make it difficult for policymakers to decipher strategies for how to use online learning.

Some of the greatest debates surrounding online education seem to be a battle of the potential of online learning versus the actual outcomes. The potential of online courses includes students having 24/7 access to materials, flexible pacing, access to interactive materials, and immediate feedback on work (Grigoryan, 2017; Heppen et al., 2017; Stevens et al., 2016; Viano, 2018; Wheatley, 2016; Yilmaz & Keser, 2017).

However, it has been strongly indicated in many studies that younger students or students who take online courses for credit recovery tend to have significantly lower general reading and math skills (Heinrich et al., 2019; Heissel, 2016; Heppen et al., 2017), and most likely have significantly lower technology skills as they probably come from backgrounds that lack technology at home (Heinrich, et al., 2019; Viano, 2018). Additionally, many online learning studies were finding that students were often distracted or spent the majority of their online time off-task or on websites other than their course (e.g., Heinrich et al., 2019; Yates et al., 2021); students found online courses to be much more difficult than their f2f counterpart (e.g., Heppen et al., 2017), and feedback in a self-paced course, and communication in general in an asynchronous course, was often delayed (e.g., Dockter, 2016; Oliver et al., 2009). Other possible conclusions of online learning were very hard to come by as the results of studies were varied depending on location, decade (1990s, 2000s, 2010s, etc.), school term(s), and framework applied to conduct the research (Alfrisa et al., 2020; Viano, 2018).

Even as online learning continues to grow at astounding rates, the true power it has for aiding ninth-grade Algebra I students is still unknown. Heppen et al.’s (2017) study provided the most comprehensive research for online versus f2f, but even this research was very narrowly focused on Algebra I credit recovery taken over a very short summer course. Other research that compared online to f2f, even if the focus was Algebra I, either did not specify if the online courses were provided by a district or state or to which population of f2f students they were being compared. What is still not known about online Algebra I courses is accurate dropout and completion numbers (Viano, 2018), and the quality of the online Algebra I courses, especially compared to their f2f counterparts (Clements, Stafford, et al., 2015; Clements, Zweig, et al., 2015; Heinrich et al., 2019), or whether students in online courses were truly learning the material or just getting by well-enough to earn course credits (Heinrich et al., 2019). This study aimed to answer some of these questions.

Well-designed research studies can provide strong, accurate evidence for creating change in policy or practice while "poorly designed studies are dangerous because of their potential to influence practice based on flawed methodology" (Kendall, 2003, p. 168). In the coming chapter, aspects of the research methodology, population, validity, and ethical concerns are discussed in detail to provide a foundation for creating a well-designed study.

Chapter 3: Research Method

The purpose of this quantitative study was to assess and compare learning outcomes for one U.S. region's students who took an Algebra I course in a district-run virtual program, a state-run virtual program, or in a f2f environment to determine the efficacy of the two approaches for moving these students closer to meeting graduation requirements. This chapter provides details as to the research design, the methodology (including population, sampling procedures, data collection procedures, variable operationalization, and data analysis), validity considerations, and ethical concerns for this study.

Research Design and Rationale

According to the WWC (2017a), a program of the U.S. Department of Education's Institute of Education Sciences (IES), "it is critical that education decision makers have access to the best evidence about the effectiveness of education products, programs, policies, and practices" (p. 1). To save educational authorities the time of combing through all educational research studies conducted, the WWC uses a strict set of standards to identify the most accurate, credible, quality research available and summarizes the collection of evidence into a single report. To meet eligibility standards for the WWC, studies must have used either a randomized controlled trial, quasi-experimental design, regression discontinuity design, or single-case design (WWC, 2017a). To provide a study that can be considered as advancing knowledge in the field of education, the research design and methodology of my study attempted to meet all standards as described in the WWC *Standards Handbook Version 4.0* .

To suggest causation in any form, in a true-experimental design, there must be a predetermined association between the independent and dependent variables; sequencing must follow such that the change in the independent variable happened before the change in the dependent variable; and there must be substantial evidence that extraneous variables did not cause the change in the dependent variable (Coalition for Evidence-Based Policy, 2010, 2014a, 2014b; Leatherdale, 2019). Additionally, there must be random assignment of subjects to each group of the independent variable to reduce the impact of extraneous variables and allow the assumption that the control and experimental groups are the same at the start of the treatment. Unfortunately, in educational settings such as those in my study, it could not be ruled out that there were no outside influences that affected the study habits, progress, and success of students disallowing for any type of causal inferences to be made. However, quasi-experimental designs may be used to evaluate the impact of an intervention (Coalition for Evidence-Based Policy, 2014b; Leatherdale, 2019; White & Sabarwal, 2014), such as the impact of online learning on the success of high school Algebra I students. Unfortunately, with any quasi-experimental design, equivalent groups cannot be assumed. As noted earlier, data by which the groups could be calculated and assumed equivalent, such as end-of-course pre-algebra grades, were not available, and the use of growth scores could help mitigate the differences in the groups at baseline.

I conducted this quantitative study using a nonequivalent group quasi-experimental design. More specifically, quasi-experimental research designs compare control and treatment groups that are shown to be as similar as possible with the differing characteristic being the intervention or program being evaluated (WWC, 2017a), and assignment to the groups is conducted through self-selection or administration placement rather than randomization (White & Sabarwal, 2014). For this study, the control and treatment groups were created prior to the start of the study, and all collected data was archival. The control and treatment groups were comprised of high school students who took Algebra I for the first time as ninth graders. These students were placed into either the control or treatment groups based on student choice or school administration assignment, and it was assumed that the groups were similar in prior knowledge, ability, access to technology, etc. The differing factor between the three groups were that the treatment groups took the Algebra I courses online, either through the district-run or state-run program, while the control group took the course in the traditional f2f environment. Unfortunately, after receiving the data, it was impossible to compensate for inequalities between groups. This is discussed further in Chapter 4.

This nonequivalent group quasi-experimental design was used as a means of discovering the extent of the differences that instructional environments have on student success and course completion for high school Algebra I. The independent variable included three categories: the district-run online learning environment (treatment group), the state-run online learning environment (treatment group), and the f2f learning environment (control group). In most schools across the United States, students advance onto higher grades and reach graduation based on their end-of-course grades. For this reason, the dependent variable of student success was determined based on end-of-course final grades and end-of-course state assessment scores. At the time of the study, students had to pass four math courses with a minimum course grade of 70% to graduate (GADOE, 2014, 2022b). The end-of-course state assessments counted toward 20% of students’ overall course grades. The dependent variable of course completion was measured by students' end-of-course grade codes with a grade code of A, B, C, or F indicating the student had completed the course, even if they did not successfully pass it, while a grade code of I or Z (withdraw) indicated the student did not complete the course. Student gender, ethnicity, and race served as covariates and were used to establish baseline similarity as far as demographics between the online and f2f groups. While disadvantaged status and any type of entry-level data would have further provided details for establishing baseline similarities, that information was not available at the time, as further explained in Chapter 5. To capture whether the differences between group outcomes were due to the difference in learning environments, baseline characteristics between the control and treatment groups should be as similar as possible (WWC, 2017a; White & Sabarwal, 2014). Group comparability was made to the degree possible, but with the lack of entry-level data, it must be acknowledged that this study has major threats to internal validity.

To thwart any researcher bias and to prevent contamination of data, both the online and f2f comparison groups were formed before the start of the study, and all data were obtained from the school board office archives. This strategy of data collection ensures that the researcher had no direct contact with any student or teacher; however, it did pose a possible time constraint. Since all data was supplied by the school board office, I had no control over how quickly data was collected from their archives and provided to me for analysis. Because the courses were already completed by the time data was requested, the data was readily accessible. Once collected, data were analyzed using IBM SPSS software, version 27. For RQ1 and RQ2, ordinal data was intended to be analyzed through ordinal logistic regressions while RQ3's nominal data was intended to be analyzed using multinominal regression.

Methodology

According to Kendall (2003), "Writing a thorough and comprehensive protocol in the planning stage of the research project is essential" (p. 165). The planning stage for the methodology of a study should include specifics about the population, sampling strategies, data collection, and analysis procedures. Details should be thorough enough to make the study replicable for other researchers (Kendall, 2003). This section provides such information.

Population

The target population for this study was high school Algebra I students who took their course for the first time as a ninth grader in one of three modalities: either (a) entirely online with the district, (b) entirely online using the state-run online program, or (c) f2f classroom during a regular school day. Over the school years 2016-17 to 2019-20, there were an average of approximately 20,978 students per year in Grades 9 to 12 in the study's school district (GADOE, 2020f). During the spring semester of the 2019-2020 school year, the high school student body was approximately 31% White, 40% Black, 15% Hispanic, 12% Asian, or 2% of two or more races (GADOE, 2020g). Overall, in the district during the 2020 school year, the student population for preschool to 12th-grade was approximately 33% physically or learning disabled (GADOE, 2020h). I assumed that these percentages had been similar for the school years included in this study.

Using the 2014-2017 Algebra I EOC assessment fail rates of 26.1% for the given district and assuming all ninth-grade students would take Algebra I while no students in higher grades would take the course, I estimated that an average of 1,369 students per year would fail their Algebra I end-of-course state assessment. However, not all students who fail the end-of-course state assessment fail the actual course.

Sampling and Sampling Procedures

This research study took place after students were placed in online or f2f environments, and courses were completed. It was unclear how students were placed into courses, whether it be due to student preference, scheduling conflicts, health reasons (homebound students), administrative placement (student expelled or otherwise not allowed to attend f2f classroom), or if students were randomly placed into either the online or f2f classroom. This study requested data for the total population of ninth-grade high school Algebra I students in the district for the school years specified below, making this study an example of total population or census sampling. Although not placed by chance, the sample included the entire target population and was well-defined, allowing for meaningful inferences to be made and for results to be generalized to similar populations (see Lesko et al., 2017).

The study’s sample included only ninth grade high school students who took Algebra I for the first time in the school district either f2f or in one of the two online programs, the district-run or the state-run. Finally, data were included for the school years

2014-15 to 2018-19. These inclusions and exclusions of the target population were chosen to help ensure baseline comparisons of the treatment and control groups were more equivalent as they helped to eliminate possible confounding factors (WWC, 2017a*)*. However, entry-level data was not available to ensure group comparability at the start of the study.

Even though this study surpassed the required total sample size, due to the use of census sampling, the issue of sample size was not a concern as there would be no attempt to make inferences about a population. However, when considering state virtual separate from district virtual, the sample was quite limited. A discussion of logistic regression approaches to sample sizes is provided to show that this study does surpass the required total sample size, even though the subgroup of state virtual did not.

Recent research regarding the calculations for determining sample size for ordinal or multinomial logistic regression for fields other than clinical seemed to be near nonexistent. Most logistic regression sample size research included detailed equations for calculating sample size, specifically for binomial or multilevel logistic regression, and varied greatly from one article to the next (see Adhikari, 2021; Ali et al., 2019; Bartlett et al., 2001; Bush, 2015; Li, 2014; Olvera Astivia et al., 2019; Peduzzi et al., 1996; van Smeden et al., 2016, 2019). Though equations differed, most researchers agreed that under sampling could cause problems in the accuracy of logistic regression calculations (e.g., Adhikari, 2021; Ali et al., 2019; Bartlett et al., 2001; Bush, 2015; Li, 2014; Mascha & Vetter, 2018; Olvera Astivia et al., 2019; van Smeden et al., 2019), including causing separation of likelihood (Santos & Barrios, 2017; van Smeden et al., 2016), inaccurate sample variance and regression coefficient bias (Peduzzi et al., 1996; Timberlake, 2011; van Smeden et al., 2016), or low prediction accuracy (Adhikari, 2021; Timberlake, 2011; van Smeden et al., 2019). Studies that did not report specific equations often suggested sample size be calculated using a minimum events per variable (EPV) ratio (see Grant et al., 2019; Hair et al., 1995; Halinski & Feldt, 1970; Miller & Kunce, 1973; Peduzzi et al., 1996; van Smeden et al., 2016). Some researchers suggested the number of EPV should certainly not drop below five (see Ahmad & Halim, 2017; Hair et al., 1995), and that logistic regression calculation problems are minimized or nonexistent when EPV is 10 or greater (Grant et al., 2019; Halinski & Feldt, 1970; Miller & Kunce, 1973; Peduzzi et al., 1996). Other researchers recommended that EPV should be between 10 to 20 (Austin & Steyerberg, 2017; Harrell et al., 1985), some argued the EPV ratio should be at least 30 (Dhivyadeepa, 2015; Pedhazur & Schmelkin, 1991), and still others suggested that total sample sizes for regression models should not be less than 100 participants to maintain statistical power (Long, 1997; Maas & Hox, 2005). However, Ahmad and Halim (2017) argued that "raising the sample size above the level indicated by the sample size formula will increase the type I error" (p. 30). Additionally, Ahmad and Halim suggested that sampling should have been used instead of population census data in most cases, but this was mostly due to low response rates. I used census sampling, but data was received from the school district archives so students did not have a choice in not responding.

My study had one independent variable with three categories: f2f, district virtual, and state virtual. If the number of EPVs were to follow one of the larger recommendations of 30 participants per variable, the minimum total sample size number would be set to 30. My total sample size met or exceeded all the above-recommended ratios, as well as the minimum total sample size of 100 suggested by Long, (1997) and Maas and Hox (2005). Being that the sampling procedure used census sampling, the resulting sample was much larger than 30 students. The total sample size for this study was 26,887 participants, with 26,563 students in the f2f environment, 307 students in the virtual district environment, and 17 students in the state virtual environment.

Procedures for Recruitment, Participation, and Data Collection

Due to the use of archived data from previous students, recruitment and participation of current students were not a concern for this study. I followed the Walden University Institutional Review Board (IRB) and cooperating school district's approval processes for the collection and use of archived data, as described below. Once IRB approval was received (IRB approval No. 03-16-20-0381762), archived data was requested to be supplied by the cooperating school district. The request specifically asked that the data collected be redacted so there was no personally identifying information for students and that it only include gender, ethnicity, race, socioeconomic status, end-of-term Algebra I course grades, and most recent Algebra I EOC assessment scores.

Operationalization of Constructs

For this study, the variables to be assessed were defined primarily in accordance with the International Association for K-12 Online Learning's (iNACOL) report, *The Online Learning Definitions Project* (2011), with the State DOE and the cooperating school district's definitions being used as needed. The variables for this study included the independent variable of Algebra I course environment, along with the dependent variables of student course success, student Milestone assessment success, and course completion. The confounding variables were gender, ethnicity, and race.

Algebra I Course Environments

The Algebra I course environment is the educational setting through which a student was taking their course, either online or in a f2f classroom. The online learning environment was when "instruction and content are delivered primarily over the internet" (iNACOL, 2011, p. 7). The online environments of this study included both the district-run online Algebra I courses and the state-run online Algebra I courses. Students who took the Algebra I course in the f2f environment were those who were expected to attend a regularly scheduled class at the school and meet in person with their instructor(s). Additionally, it was assumed that the f2f, district virtual, and state virtual courses all conformed to state standards and were therefore similar in content.

For this study, the variable for learning environment was based on the school district's coding system (either through course code or online course indicator as mentioned in the definitions) as to whether the student took the course in one of the online environments or the f2f environment. The learning environment was measured as a nominal variable dummy coded with the f2f environment coded as 1, the district-run online environment coded as 2, and the state-run online environment coded as 3. For example, when running the regression calculations in SPSS, students who took the Algebra I course in the f2f environment had an independent variable value of 1, while students who took the course in the district-run online environment had an independent variable value of 2.

Student Course Success

Student course success was measured by considering student course grades. Students who had remained in a course until the end-of-the term received a course grade of A, B, C, or F (D was not included on this grading scale). Students were considered to have passed a course if they received a grade of 70% or higher while students who received an F (69% or lower) were considered to have failed the course. Each local school board could determine how letter grades were distributed so long as the minimum passing score for a C was 70% (GADOE, 2014). Grades for the participating district were based on a 10-point scale with 90-100 an A, 80-89 a B, 70-79 a C, and 69 and below an F. Course grades were measured as ordinal variables with grades of A receiving a rank of 1, Bs a rank of 2, Cs a rank of 3, and Fs a rank of 4. For example, if Student A earned a grade of 84%, B, their course grade value in SPSS would be 2.

Student Standardized Assessment Success

The original plan proposed the use of the state assessments, known as the Milestones, to be used as an additional measure for determining overall student success for both RQ1 and RQ2. The plan was to use this assessment because, while passing the state assessment for Algebra I was not technically a requirement for successfully completing a course, it did weigh into a student's overall course grade. The state assessments counted toward 20% of a student's final course grade, and the final course grade had to be 70% or higher to pass (GADOE, 2014, 2020b). This meant failing a state assessment could potentially cause a student to fail the course.

The data received did not include the Milestones EOC assessment scores. Instead, FastBridge Growth Assessment scores were included, but only for a small number of students. The lack of data for this variable limited the ability to conduct in-depth analysis, other than to provide descriptive statistics. More detail is provided in Chapter 4.

FastBridge assessments are screening tools given as assessments to all students in the district during three testing timeframes throughout the year, fall, winter, and spring (Illuminate Education, 2022a). The assessments are used as progress checkpoints to identify students who may be at high-risk, medium-risk, and low-risk of not meeting benchmark standards and not completing a course, with low-risk students being more likely to graduate and attend college (Biery, 2022; Illuminate Education, 2022a). AMath assessments contain approximately 30 questions that adapt in difficulty depending on a student's performance on prior items (Illuminate Education, 2022b). National norms from the fall of 2019 resulted in a scaled score range between 150 and 250, M = 224.2, and SD = 12.3 for Grade 8 aMath (Illuminate Education, 2022c). National norms for ninth to 12^{th}-grade were not yet available.

Course Completion

For this study, the original plan was to include students in the course completion calculation if they were enrolled in the course after the ten-day add/drop period (GADOE, 2011), they remained in the course until the end of the fixed semester, and they received a final course grade of A, B, C, or F (Los Angeles Mission College, Office of Institutional Research and Planning, 2011; The Research and Planning Group for California Community Colleges, 2011). The state DOE has set a series of over 25 different withdrawal codes for students who drop out of a course, transfer to a new school, are expelled, etc. (GADOE, 2020a). The cooperating district used only the withdrawal code of "I" rather than the multitude of codes that indicated students were withdrawn due to other reasons, such as court order (code C) or a transfer to another school (for example, J, K, N, etc.). For the purposes of this study and due to the lack of variation in the withdrawal codes utilized by the district, all students who were withdrawn after the ten-day add/drop period, regardless of the reason, would have been included in the analysis for RQ3. Students who received a grade, indicating completion, were dummy coded as 2 while students whose term grade was coded I were coded as 1, indicating a withdrawal from the course.

As with the state assessment scores, the number of students who had withdrawn was too small for analysis other than descriptive statistics (only 17 for the virtual courses combined). More detail on this is provided in Chapter 4.

Gender and Ethnicity

The covariates of student gender, ethnicity, and race were initially planned to be used separately as a means of determining baseline equivalence between the two learning environments and included only values used in the state DOE reporting standards. Possible values for gender included only male or female and were coded 1 for males and 2 for females.

The state DOE (2020a) defined ethnicity as being either Hispanic or non-Hispanic, while race included identifying as White, Black/African American, Asian, Multiracial, American Indian or Native Alaskan, or Native Hawaiian or Pacific Islander.

The provided data set seemed to be a combination of both ethnicity and race and was therefore treated as one variable of ethnicity/race rather than two separate variables. In the provided data set, the title heading was federal *ethnicity* code, so in this study I labeled any further mentions of ethnicity or race as just ethnicity. The available options for ethnicity were White, Black/African American, Hispanic, Asian, Multiracial, American Indian or Native Alaskan, or Native Hawaiian or Pacific Islander. The coding for this variable was intended to be adjusted so that Black/African American = 1, White = 2, Hispanic = 3, Asian = 4, Multiracial = 5, and American Indian or Native Alaskan and Native Hawaiian or Pacific Islander would be combined to be coded as 6. As an example of this coding strategy, consider a student who was a Black female. This student would have received a code of 2 for gender and 1 for ethnicity/race. More detail on this is provided in Chapter 4.

Data Analysis Plan

For this study, data was collected and analyzed in an attempt to answer the following research questions.

RQ1: Is there a difference in student course success, as measured by end-of-course grades, between local online courses, state-run online courses, and f2f instructional environments in Algebra I courses, while controlling for student demographics, such as gender, ethnicity, and race?

$H_0$1: There is no difference in student course success between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

H_a1: There is a difference in student course success between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

RQ2: Is there a difference in student state assessment success, as measured by SOL assessment scores, between local online courses, state-run online courses, and f2f instructional environments in Algebra I courses, while controlling for student demographics, such as gender, ethnicity, and race?

H_02: There is no difference in student course success between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

H_a2: There is a difference in student course success between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

RQ3: Is there a difference in course completion (as measured by course grade codes) between local online courses, state-run online courses, and f2f instructional environments, while controlling for student demographics, such as gender, ethnicity, and race?

H_03: There is no difference in course completion between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

*H*a3: There is a difference in course completion between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

As noted in Chapter 1, research suggested there were differences in student academic achievement based on geographic location, disadvantaged status, ethnicity, race, gender, and student age/grade level (e.g., Cavanaugh & Jacquemin, 2015; Chapman et al., 2011; Cooper et al., 2019; Curtis & Werth, 2015; Heinrich et al., 2019; Hughes et al., 2015; Protopsaltis & Baum, 2019). Therefore, student characteristics were also considered when comparing at-risk students (Conway et al., 2016).

Once collected, data were stored and analyzed using SPSS software. SPSS is a software program designed to aid researchers in storing, coding, and analyzing data with a wide array of statistical calculations. SPSS was used to run logistic regression calculations for each of the research questions. For RQ1, ordinal logistic regression was used because the dependent variables consisted of ordinal values of A, B, C, or F. Ordinal logistic regression was intended to be used for RQ2 because Illuminate Education's *FastBridge aMath Score Interpretation Guide* (2019) separated raw scores into ordinal categories. However, a lack of data for online FastBridge scores prevented ordinal logistic regressions from being utilized, and instead, analysis proceeded with multinominal logistic regression for f2f data only. Multinomial logistic regression was intended to be used for RQ3, however, limitations in the data received prevented this analysis to take place. For RQ3, descriptive statistics were provided where possible. More details are provided in Chapter 4.

Before running any analysis procedures, data were screened for errors and cleaned up as needed. Hellerstein (2008) noted four sources of data error and suggested techniques for cleaning them. These four errors included data entry errors, measurement errors, distillation errors (simplifying or consolidating data prior to entry), and data integration errors (differing data collected, measured, or entered into databases over time). While the techniques for data entry at the proposed district have been somewhat streamlined to include easier interfaces for teachers to enter grades and export them to the school district's database, there was still a possibility of human error when entering information. Due to the nature of using archival data, there was no way to control for or fix data entry errors, missing values, or outliers, so data must be accepted as being entered correctly. Unfortunately, these data points were eliminated from the sample as there was no means of correcting them. The available sample was large enough that the loss of a few records did not skew the overall results so there was no need for imputation of missing data. However, in some cases of missing or incorrectly input data, some values were still of use for certain calculations. For example, if a letter were entered other than A, B, C, F, or I for the Algebra I course grade then that entry was not useful in the calculations for either student course success or course completion but could still have been utilized for analyzing standardized assessment success. Data that could be cleaned up included making sure student identifying codes were unique to each student and that there were no duplicate student identifiers without violating student identification concealment agreements.

Data screening techniques for the ordinal logistic regression calculations consisted of testing data for model fit, including testing for multicollinearity between the covariates and the independent variable and for proportional odds (Laerd Statistics, 2015). If there was multicollinearity between the independent variables, the decision was made to either continue or simply to drop one of the parallel variables from the calculation. If the assumption of proportional odds was violated, data was then analyzed as a multinomial logistic regression (Laerd Statistics, 2015). Multinomial logistic regressions did not allow for the ranking of values to hold but still allowed for categorical variables to be calculated and compared.

According to Laerd Statistics (2018), in multinomial logistic regression analysis, preliminary testing should be done to assure that there is (a) independence of observations, (b) that the dependent variable categories are mutually exclusive, (c) that there must be "a linear relationship between the continuous independent variables and the logit transformation of the dependent variable" (para. *Assumption #5*), (d) there is no multicollinearity, and (e) there are no significant outliers. Any violations of the assumptions of the multinomial logistic regression model required data to be transformed, for variables or variable values to be dropped, or for the researcher to ignore the violation and carry on (Laerd Statistics, 2018).

A major internal validity factor can be sampling bias because the difference in group sizes, demographics, prior knowledge, ability, etc., can affect regression analysis outcomes. However, due to archival data, census sampling, and multiple school years being included, sampling bias was mitigated to some degree. As previously noted, without entry-level data being available and not being able to randomly place students into learning environments, the findings of this study must be interpreted with caution.

Assuming there were no assumptions violated for either logistic regression procedure, data could then be run through calculations in SPSS. In addition to frequency tables for descriptive statistics, SPSS was used to produce several tables requiring assessments before final determinations can be made as to which, if any, independent variables significantly affected the dependent variables, and how well the logistic regression models predicted the dependent variables. These assessments included covariate patterns and cell size (zero frequencies), overall goodness-of-fit, overall parameter estimates, probability estimates, and model fit (Laerd Statistics, 2015).

The first consideration was to understand the number of covariate patterns (unique independent variable combinations) and cell size (the product of covariate patterns and categories of the dependent variable less the product of the number of unique combinations of the independent variables and the categories of the dependent variable). Cell size is the calculation of how many cells have zero frequency. Knowing this value helps in determining if the goodness-of-fit models should be approached with caution or considered reliable. Ideally, there would be no cells with zero frequency, and at least 80% of expected cell frequencies to ascertain that the goodness-of-fit measures were reliable (Laerd Statistics, 2015).

Overall goodness-of-fit tests provided measures as to how well the logistic regression models fit the data (Laerd Statistics, 2015). These measures were the Pearson and deviance chi-square statistics which showed how poorly the model fit the data.

Therefore, nonstatistically significant chi-square values would indicate a good model fit, while statistically significant values would indicate a poor model fit (Laerd Statistics, 2015). Additionally, a likelihood ratio test was generated to determine the difference between the intercept-only model and the full model. Statistically significant differences would indicate that at least one independent variable/covariate meaningfully explained the dependent variable (Laerd Statistics, 2015).

A statistically significant model fit showed that the independent variables could be used to predict the dependent variable, but not which independent variables. A parameter estimates table was used to determine which of the independent variables had a statistically significant effects on the dependent variable (Laerd Statistics, 2015). This determination was based on the odds ratios between the online and f2f environments. The odds ratios, significance, and 95% Wald confidence intervals, along with the Wald chi-square values and significance of the hypothesis test, were reported for independent variables that predicted the dependent variable (Laerd Statistics, 2015).

As a final measure of the logistic regression procedures, based on the outcomes of the probability estimates, a crosstabulation table was created as needed. If the probability estimates indicated the same outcomes as the observed dependent variable categories, then it could be assumed the logistic regression models fit the data well (Laerd Statistics, 2015). If the probability estimates of the given model incorrectly predicted the correct category of the dependent variable, then crosstabulations were conducted to determine how many cases of the dependent variable were correctly and incorrectly predicted by the model (Laerd Statistics, 2015).

Finally, results found from the regression analyses were interpreted in terms of transactional distance, to the greatest degree possible. As noted in Chapter 2, prior research has indicated that students in online courses experienced higher levels of transactional distance than f2f students and tended to underperform compared to their f2f counterparts. Should this be the case with this study, results were discussed regarding possible levels of transactional distance between the learning environments.

Threats to Validity

The ultimate goal of most research studies is to fully and accurately understand some phenomena and apply the findings to other similar situations. The perfect quantitative research study is one that proves to have high internal, external, construct, and statistical conclusion validity. Unfortunately, no study is perfect. Instead, every scientific research study runs the risk of not being completely valid. Carefully designed studies account for elements, such as confounding variables and heterogeneous populations, to create higher validity. Having high internal and external validity means study results confidently detail the relationships between dependent and independent variables, and that the results are generalizable to other populations, respectively. Construct validity signifies that the measurement techniques used within a study truly measure what they were intended to (Furr & Heuckeroth, 2019; Guerda, 2020; Laerd Dissertation, 2012, *Construct Validity*), while statistical conclusion validity considers the accuracy of statistical research findings (Garcia-Perez, 2012; Guerda, 2020). Details as to the precautions taken in the study to reduce the threats to validity are described in detail in this section.

Internal validity can be threatened due to compromised protocols such as unaccounted for confounding variables, sampling bias, participant dropout, pretesting participants, or changes in the dependent variable during research testing (Behi & Nolan, 2014; Cuncic, 2019; Glen, 2014; Lesko et al., 2017; Michael, n.d.). In this study, the covariates of students' gender, ethnicity, and race were considered and controlled for when determining the significance of the independent variable. Additionally, due to every student having an equally likely chance of being selected for the study (all students had a 100% chance) and the students being placed in the appropriate learning environment prior to the researcher's involvement, the threat to population representation and selection bias was lessened as a threat to internal validity. While internal validity can't be eliminated, the use of longitudinal data can strengthen both the internal and external validity of a study (Fruehwirth, et al., 2021; Smith Jaggars et al., 2013). Longitudinal data "directly addresses differential selection…which would affect the internal validity of [other] designs…[and] permits us to investigate underlying causes after accounting for key confounds" (Fruehwirth et al., 2021, *Introduction* para. 6). The use of longitudinal data in this study helped to alleviate confounding variables such as student demographics, student motivation and readiness, teacher biases, and policy changes over the years. While longitudinal data, such as that used in this study, could reduce internal and external validity threats, it is still advised to interpret results with caution.

This study's data was in the form of archived information regarding students' grades and demographics. Hageman (2017) stressed that the use of archival data can compromise internal validity due to its inability to associate the causality of the independent variable with the dependent variable. However, Hageman (2017) also suggested that one of the ways of increasing internal validity when using archived data is to investigate “the relationship between a naturally occurring event and a comparison event” (p. 8), such as in educational settings, when variables cannot be easily manipulated. For this study, the archived data utilized were the results of Algebra I courses that took place in a natural educational setting. Additionally, the use of archived data means that there was no threat of participant dropout (in terms of participants leaving the study, not dropping out of the course), Hawthorne effect (participants changing their behavior due to knowing they are being observed), changes in dependent variables (for example, participant maturation), or inaccuracies due to pretesting (see Cuncic, 2019; Garcia-Perez, 2012; Glen, 2014; Salkind, 2010). Although elements exist that could potentially still influence study results, using archived data from a natural setting and accounting for confounding variables aided in increasing internal validity (Cuncic, 2019; Garcia-Perez, 2012; Glen, 2014; Hageman, 2017; Salkind, 2010).

Finally, for this study, the sample was comprised of the entire population over several years. The population was clearly defined as first-time Algebra I students who took their course in a natural educational setting either online, through the district-run or state-run programs, or f2f during their ninth-grade year in the cooperating school district and were not considered to be in honors or remedial Algebra I courses. Due to the study utilizing archived data, there was no risk of students knowing their results were used in a study that may have altered their behavior or test performance. This study was not replicated, therefore, comparing the results to findings in other studies was not possible.

External validity addresses the ability of study results to be generalized to other populations, settings, and times (Cuncic, 2019; Garcia-Perez, 2012; Glen, 2015a; Lesko et al., 2017; Michael, n.d.). Often, external validity is believed to be higher when the study sample is randomly selected from a well-defined target population (Cuncic, 2019; Lesko et al., 2017; Michael, n.d.). According to Hageman (2017), the use of archival data can increase external validity due to its usefulness in examining large amounts of data for trends that influence multiple populations or settings. Cuncic (2019) suggested there are ways to increase external validity, even if the sample is not randomly selected. These practices include clearly defining the target population, using natural settings, making sure participants feel they are not being studied, repeating the study with different samples or settings and comparing results, and calibrating data as necessary using statistical methods (Cuncic, 2019). However, since this study used census sampling, these threats to external validity did not apply.

The constructs of a research study are the mental concepts upon which a study is based and measured by selected variables (Garcia-Perez, 2012; Guerda, 2020; Laerd Dissertation, 2012). Construct validity is highly dependent on how well the operational definitions of the constructs and variables of a study are defined (Garcia-Perez, 2012; Guerda, 2020; Laerd Dissertation, 2012). Construct validity can be threatened due to poorly defined constructs, insufficient variable measurements, insufficient measurement methods, participant sensitivity to the study treatment, researcher expectations, or when constructs have an overlapping relationship or meaning (Guerda, 2020).

For my study, operational definitions were mainly based on those set by iNACOL, one of the leading authorities on educational research and advancing educational reform (iNACOL, 2019; Schwartz, 2018). Using the definitions by iNACOL helps to ensure that the constructs were well defined and lack the possibility of having multiple or overlapping meanings, thus increasing construct validity. Additionally, construct validity is heightened due to the use of archived data and courses being completed prior to the start of the study making them unaware of being part of a study.

The greatest threat to construct validity for this study was the lack of sufficient measurement tools and procedures for each of the variables. The means of measuring student success was by comparing students' Algebra I course grades and end-of-course state assessment scores between the online and f2f environments. Other measurements, such as teacher surveys or measurements of course rigor, could prove to be useful in truly understanding how well students performed in the respective courses but were beyond the scope of this study. Course completion was measured by whether a student who enrolled in the course completed it to the point of earning a course grade, even if that grade was an F. It would be helpful to know whether students who did not complete a course did so because they chose not to participate, were incarcerated, were noncompleters due to medical reasons, etc. While this information would not change the percentage of students who were not completing Algebra I in the online and f2f environments, it would give a clearer look as to why students were not completing and could allow researchers to move toward finding ways to keep more students successfully progressing through the courses.

Statistical conclusion validity (SCV) refers to whether the collected data in a research study were properly analyzed to confidently report on the relationships between dependent and independent variables while maintaining awareness of Type I and Type II error rates (Garcia-Perez, 2012; Guerda, 2020). Common threats to SCV include mining the data (fishing for results), low statistical power, improper treatment execution, using unreliable measures, or violating statistical test assumptions (Glen, 2015b; Guerda, 2020).

Data mining involves combing through large data sets or repeating statistical tests to find desired trends or results (Glen, 2015b; Guerda, 2020), often affecting the Type I error rate (Garcia-Perez, 2012; Guerda, 2020). The data collected for the study was used in full, (other than being filtered to include only the identified population), not sorted through looking for trends, and was only run through the ordinal or multinominal logistic regressions once per research question. Due to the use of census sampling, the SCV threat of having low statistical power was not an issue. Before running the ordinal or multinomial logistic regression procedures, data were tested to ensure assumptions were met (as discussed above) to alleviate threats to SCV due to violations of test assumptions. Running both a power analysis and procedures to test assumptions aided in increasing the validity of statistical conclusions.

As for treatment fidelity, data was collected after the Algebra I courses had been completed (for those who finished the course), so implementation of the “treatment” (online versus f2f courses) was performed by the respective course teachers prior to data collection. While it was expected that there would be some discrepancies between the intended curriculum and what students actually learned, in both f2f and online courses, an unfortunate real-world expectation, it was assumed that these differences would be minimal. It was beyond the scope of this study to delve into treatment fidelity considerations, but an acknowledgment that issues existed was a consideration when interpreting the results of this study. Due to the courses being completed in advance of the collection of data, teachers and students could not have known about the study, so normal behaviors were not compromised, and SCV was not threatened as a result of improper treatment execution.

The greatest threat to SCV was from the measure of student end-of-course grades due to the subjectivity of grades imparted by the teachers. Meissel et al. (2017) found that there were significant differences in New Zealand teachers' judgments of student achievement based on students' characteristics (male, female, indigenous New Zealanders, and Pacific Islanders). Considering Meissel et al.'s (2017) study, there was some concern that students' grades in this study could have been influenced by teacher judgments rather than by true abilities, posing a possible threat to SCV. Unfortunately, it was beyond the capacity of this study to be able to assess whether teacher judgments affected students' grades, so, for this study, it was unknown if this assumption was true or not.

While the eventual objective of any research study is to expand the knowledge of a given field by providing sound evidence that is indisputable, it is never the case. In one way or another, the validity of a study is threatened. It is up to the researcher(s) of the study to understand where threats may lie and create solutions for minimizing them. In this study, the greatest threats to validity came from the use of the chosen measurements and the inability to account for such things as teacher bias when grading, inaccurate enrollment counts, and a lack of multiple measurements for each variable. However, the use of a census sample of longitudinal archived data can help ensure some validity internally, externally, and statistically. Archived data obtained from a natural setting can aid in eliminating threats due to sampling bias, participant dropout, and the Hawthorne effect. Further, internal validity is improved by considering extraneous variables that could influence the relationship between the dependent and independent variables, while threats to external and construct validity were reduced on account of having an unambiguously identified target population and operational definitions established by iNACOL. However, this study was limited in its ability to confirm validity due to eliminating the threat of improper treatment implementation, such as teacher effects. It was beyond the scope of this study to consider the conditions under which any differences between the learning environments may have occurred.

Ethical Procedures

When conducting research studies, researchers must consider where ethical issues may exist, such as when collecting data from their own workplace, involving minors, or requesting data regarding sensitive information. For this study, data collection procedures, requested information, and participants were aligned to follow the guidelines set forth by Walden University's Center for Research Quality (CRQ). This section included a discussion as to the procedures for ensuring ethical parameters were met, such as the necessary agreements needed to gain access to data, the ethical concerns related to data collection and storage, as well as participant protections.

As noted previously, this study utilized archived data collected from the cooperating school board office. Before obtaining data, however, permission to request data from the participant school district had to be granted by Walden University's IRB, then access to the data had to be granted by the cooperating school district's research review committee. Access was only granted after I completed the district's specific research request form and included all additional information as required.

The use of archived data is recommended by the CRQ as it "is the most ethical way to study [participants]…because it does not ask them to do anything out of the ordinary for research purposes" (Walden University, CRQ, 2019a, *General tips for avoiding ethical problems in doctoral research*). Additionally, the use of secondary or archival data allows for collection to take place without interrupting an educational or workplace setting, and for experiment and control groups to be considered without the threat of needing spare time or energy to create groups as the groups had already been created. The use of archived data in this study was also advantageous in that the intervention under consideration had already been applied to the educational setting. This meant no further steps needed to be taken to ensure the treatment was offered to the control group or that incentives needed to be used to attract participants.

Digitally archived data, such as the data that was requested for this study, allowed for simpler strategies in protecting participant information and rights because it could more easily redacted and encrypted to hide information that was personally identifying.

Redacted information included student names and school identification numbers, as well as phone numbers, addresses, and social security numbers as applicable to protect participants.

Data was requested to be sent on an encrypted USB drive or encrypted email, whichever was preferred by the school district. Once the email had been collected, data was stored on a password-protected personal cloud device (WD OneDrive) accessible only through a personal webpage and multi-step verification process. Data was only accessible by the researcher or dissertation committee, as necessary. Additionally, data will be held for 5 years for possible study replication, after which it will be destroyed.

The greatest ethical concern for this study was that because the data was regarding mostly minors, there was a question of whether parent consent was needed. According to Walden University CRQ's (2019b) frequently asked questions website, parent consent would not be needed if the results of the study could be used by the school district to directly benefit the student, the information collected was standard practice, participant information could be kept anonymous, and that the principal or school district was willing to sign a Data Use Agreement releasing de-identified data for research purposes. For this study, (a) one of the main purposes was to better understand how ninth-grade Algebra I students fair in varying environments so administrators can make more informed decisions as to student placement in online or f2f courses, (b) the data requested included student grades and Milestone assessment scores which were collected every semester as standard practice, (c) as noted earlier, personally-identifying information was redacted, and (d) the benefit of using archived data was that the data requested cannot be given unless the school district was already willing to sign a Data Use Agreement. Additionally, because the courses had already taken place and the data was archival, there were no direct or indirect interactions between the researcher and students. Based on the guidelines of the Walden University CRQ (2019a) and the specified data collection procedures, parent consent was not necessary.

Finally, concerns about ethical practice revolve around whether the individual researcher or researchers have been trained in conducting fair and ethical research. Walden University requires all students to have completed ethical research training through either the National Institute of Health (NIH) or the Collaborative Institutional Training Initiative (CITI). I completed the NIH training in 2012 (see Appendix A).

Summary

The purpose of my study was to assess and compare learning outcomes for one U.S. region's students who took an Algebra I course in a district-run virtual program, a state-run virtual program, or in a f2f environment to determine the efficacy of the two approaches for moving these students closer to meeting graduation requirements. A nonequivalent group quasi-experimental design was used to evaluate the impact of the learning environment on student success and course completion for these at-risk students. Census sampling was utilized as a means of analyzing archived data for all ninth-grade Algebra I students who took their course in the cooperating district f2f, in the district-run online program, or the state-run online program over the past 5 school years. Once data was collected, cleaned, and screened, SPSS was used to run logistic regression analyses, controlling for gender and ethnicity/race.

While no research study is perfect, it was my aim to best account for all possible threats to validity and to carefully consider all ethical issues that may exist when designing the research procedures. In this study, threats to internal and external validity were eased by considering confounding variables and using archived data collected from a naturally occurring setting, while SCV was increased by utilizing operational definitions based on those set forth by one of the leading authorities of educational research, iNACOL. While archived data can aid in eliminating threats to internal, external, and SCV, elements such as the lack of sufficient measurement tools, the subjectivity of student grades, and possible inaccuracies in enrollment numbers can still threaten construct validity and SCV.

Ethical considerations for my study included data collection and storage procedures, participant protections, and permissions needed to access data. Once permission was granted by both Walden University's IRB and the cooperating school district's research review committee, data was requested in redacted form from the school district through encrypted email or USB and stored on a password-protected personal cloud device. The greatest ethical concern for my study was in regard to the participants being minors. However, due to the type of data, student information being requested, and due to the nonexistent interactions between the researcher and students, student participants were fully protected.

Chapter 4: Results

The purpose of this quantitative study was to assess and compare learning outcomes for one U.S. region's students who took an Algebra I course in a district-run virtual program, a state-run virtual program, or in a f2f environment to determine the efficacy of virtual and f2f approaches for moving these students closer to meeting graduation requirements. To do so, data was collected and analyzed to answer the following questions:

RQ1: Is there a difference in student course success, as measured by end-of-course grades, between local online courses, state-run online courses, and f2f instructional environments in Algebra I courses, while controlling for student demographics, such as gender, ethnicity, and race?

> $H_0$1: There is no difference in student course success between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.
>
> H_a1: There is a difference in student course success between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

RQ2: Is there a difference in student state assessment success, as measured by SOL assessment scores, between local online courses, state-run online courses, and f2f instructional environments in Algebra I courses, while controlling for student demographics, such as gender, ethnicity, and race?

$H_0$2: There is no difference in student course success between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

H_a2: There is a difference in student course success between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

RQ3: Is there a difference in course completion (as measured by course grade codes) between local online courses, state-run online courses, and f2f instructional environments, while controlling for student demographics, such as gender, ethnicity, and race?

$H_0$3: There is no difference in course completion between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

H_a3: There is a difference in course completion between local online courses, state-run online courses, and f2f instructional environments when controlling for student demographics.

This chapter provides a discussion of the data that was collected, including any discrepancies from the original plan, as well as the results of statistical calculations for each of the research questions presented. Additionally, statistical calculations were interpreted as a means of answering each research question.

Data Collection

Many research studies use various sampling procedures to obtain data which must then be determined as to how closely the sample represents the population. For this study, due to census sampling, the sample was representative of the population and did not need further analysis. As noted previously, all requested data was archival and was therefore readily available as soon as requested. However, clarifications were needed to understand the abbreviations used, as well as justification for missing, substituted, and added information. The change in information of the given data set required necessary adjustments to be made due to the discrepancies from the proposed analysis strategy that was created prior to acquiring the actual data. The deviations in the provided data set included the absence of race as a variable, the addition of ninth through 12th grade credit recovery courses, the addition of flags for special education and students taking remedial courses, and the replacement of Milestone state assessment scores with FastBridge Growth data (raw scores and ROI). Data for credit recovery courses, remedial courses, and special education indicators were not used for this study; however, using this data set, could be beneficial in future studies comparing online and f2f learning.

According to the state board of education website, there is a difference in student ethnicity and student race (National Forum on Education Statistics, Race/Ethnicity Data Implementation Task Force, 2008). The state DOE defined ethnicity as being either Hispanic or non-Hispanic while race included identifying as White, Black/African American, Asian, Multiracial, American Indian or Native Alaskan, or Native Hawaiian or Pacific Islander. In the provided data set, the title heading was federal ethnicity code, and the available options were listed as the aforementioned races along with Hispanic. The provided data set seemed to be a combination of both ethnicity and race and was therefore treated as one variable of ethnicity rather than two separate variables. The coding for this variable was intended to be adjusted so that Black/African American = 1, White = 2, Hispanic = 3, Asian = 4, Multiracial = 5, and American Indian or Native Alaskan and Native Hawaiian or Pacific Islander were combined to be coded as 6. As an example of this coding strategy, consider a student who was a Black/African American female. This student would have received a code of 2 for gender and 1 for ethnicity. However, some ethnicities had to be further combined to reach minimum cell frequencies for the regression statistics. The four ethnicities with the smallest proportions, Asian, Multiracial, American Indian or Native Alaskan, and Native Hawaiian or Pacific Islander, were combined and coded 4. Additionally, to make various comparisons, coding had to be adjusted to change the reference category.

The originally planned study was to consider the differences in f2f versus virtual courses, both district and state-run, for ninth-graders taking Algebra 1 for the first time. However, there was a lack of data for the state virtual program. Among the 26,887 students taking Algebra I for the first time, only 17 represented the state-virtual program falling short of the 30 per variable necessary for accurate calculations. Due to the lack of data for this variable, calculations were completed for f2f to district virtual, and again district virtual to state virtual where possible.

Finally, data was requested regarding the raw scores of the Milestones state assessment. However, data was supplied for the FastBridge growth assessments, including raw scores for the Fall, Winter, and Spring assessments, but only for the school year 2018-2019 and only for some students. It was unclear why only the 2018-2019 data were provided, but it may be that this was the first year the school district used FastBridge rather than the Milestones assessments. Illuminate Education (2022d) described the specific assessment given as

> A simple, efficient, computer adaptive measure of both broad and component math skills from kindergarten through eighth grade. It is designed to identify students with deficits in math achievement and predict performance on state accountability measures. Used for universal screening and instructional leveling, it provides skill-based diagnostic reports of strengths and weaknesses. (*Assessment Overview*).

Using Illuminate Education's (2022e) aMath, a computer-administered adaptive screening test , students completed two or more of the same assessments within a specified time frame (Brown, 2021). FastBridge analysis measures the amount of learning growth a student has achieved within that time frame, typically from one quarter to the next. This ROI was calculated by finding the difference between 2 quarters' assessment scaled, raw scores and dividing by the number of weeks between the two assessments. For example, if a student scored a 39 on a fall assessment and a 62 on a winter assessment and there were 15 weeks between the assessments, then the ROI would be the difference between 62 and 39, which is 23, divided by 15. This would calculate as a rate of improvement of 1.533. Any ROIs resulting in a negative value indicates learning loss between the two assessments/quarters (Bielinski et al., 2020).

The dataset received from the district included raw scores for the fall, winter, and spring semesters of the 2018-2019 school year. Raw scores for only the spring were used to compare learning outcomes for the f2f and virtual programs. Raw scores for the spring semester only were chosen because they were most representative of the learning outcomes after the learning in the different environments had taken place, and the fall and winter scores were mostly for students outside of the desired population (different grades, special education, credit recovery, etc.).

Baseline Statistics

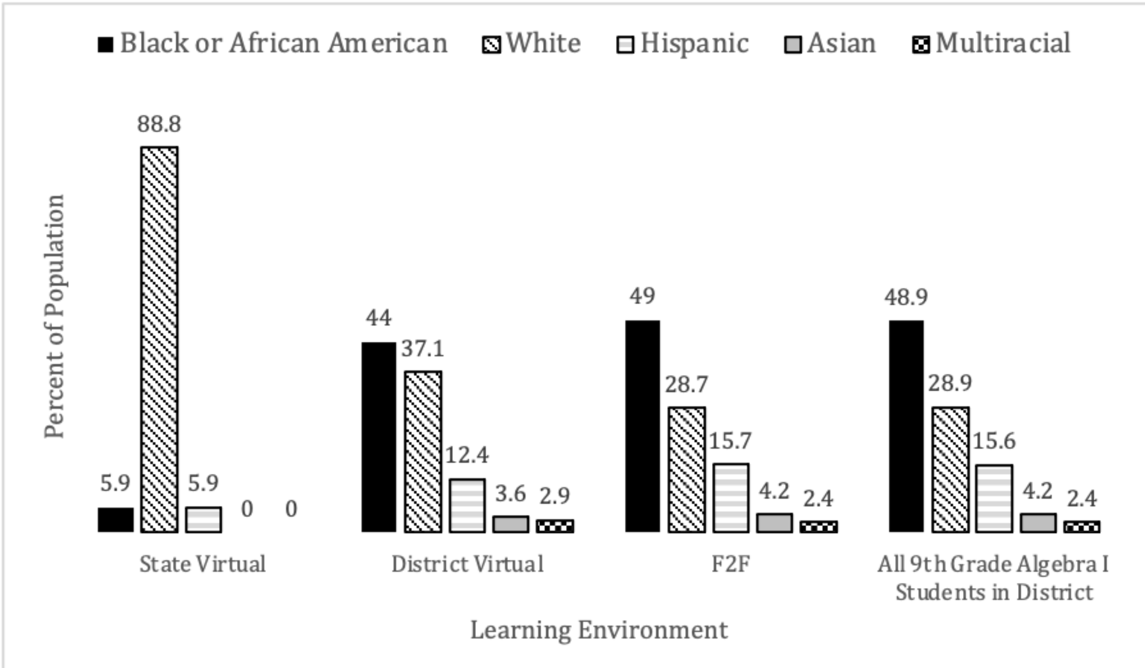
Before evaluating data between the three learning environments, I conducted baseline statical analysis to determine group similarity. A chi-square goodness-of-fit test was conducted to determine if the ethnicity distribution of the f2f and district virtual environments were proportional to that of the district's population for ninth-grade students taking Algebra I. The district's population based on ethnicity grouping consisted of 48.9% Black/African American students, 28.9% White students, 15.6% Hispanic students, 4.2% Asian students, and 2.4% Multiracial students. These percentages were used to perform chi-square goodness-of-fit tests with unequal proportions. Students taking Algebra I f2f had a minimum expected frequency of 637.5. The chi-square goodness-of-fit indicated that the population of the f2f environment was similarly distributed to the population of all district ninth-grade Algebra I course takers $\chi^2(4) = .469$, $p = .976$. These analyses were conducted for all ninth-grade Algebra I students taking the course for the first time, whether they completed the course or not.

Students taking Algebra I in the district virtual environment had a minimum expected frequency of 7.4. The chi-square goodness-of-fit indicated that the population of the district virtual environment was statistically significantly different than the distributed population of all district ninth-grade Algebra I course takers $\chi^2(4) = 11.4408$, $p < .05$. Due to census sampling of longitudinal data, even though the chi-square goodness-of-fit test for the distribution of students in the district virtual course compared to the distribution of the specified set of students in the entire district indicated the populations were not similar, regression analysis was still conducted for these two environments. However, as noted earlier, results must be considered with care.

The state virtual program had a total of 17 students from the cooperating district, 15 (88.2%) of which were White students, one Black/African American (5.9%), and one Hispanic (5.9%). Visual considerations of the population proportions indicated that populations were not similar. Figure 1 shows the population distributions for students in the state virtual program, district virtual program, f2f courses, and all ninth-grade Algebra I students in the district.

Figure 1

Population Distributions for Different Learning Environments for Ninth-Grade Algebra I



Note. Population proportions are shown prior to combining Asian and Multiracial into one category to show percentages of the five largest ethnicities. All included values were population percentages of the given environment. Population distributions included students who did not complete the course but did not include students in credit recovery courses. Student populations shown are for the school years 2016-17 to 2019-20.

After combining the ethnicities of Asian and Multiracial, the new proportions for the Asian/Multiracial category remained 0% for state virtual, 6.9% for district virtual, 6.6% for f2f, and 6.6% for all district students. After re-testing for similar distributions between the district virtual and all district Algebra I ninth graders, the minimum expected cell frequency was 19.1. The chi-square goodness-of-fit continued to indicate that the population of the district virtual environment was statistically significantly different than the distributed population of all district ninth-grade Algebra I course takers ($\chi^2(3)$ = 10.564, p = .022).

I also used the chi-square goodness-of-fit to determine if gender distributions for each of the three environments compared to the overall district population gender distribution proportions. The overall ninth-grade Algebra I population was 50.9% male and 49.1% female. Results indicated that each of the learning environments, f2f $\chi^2(1)$ = .003, p = .955, district virtual $\chi^2(1)$ = .098, p = .755, and state virtual $\chi^2(1)$ = 2.637, p = .104 were proportionately similar to the overall district.

Proportionate population distributions between the sample and given population are important in quantitative research as they allow for findings to be more accurately interpreted and for inferences about the population to be made (Ahmad & Halim, 2017). Even though the population distribution for the virtual courses did not precisely represent the population of the entire district, these subgroups were in themselves, separate populations, particularly because the virtual courses were the entire population for each group. When interpreting results, one must remember that because these individual populations were not similar, comparisons between one population to the next cannot be made. However, inferences regarding the district virtual courses may be made to future district virtual courses and students. The same may hold true for the state virtual courses as well.

Analysis for RQ1 - F2f and District Virtual

Descriptive statistics are representations of the characteristics of a given data set. In this data set, descriptive statistics were compiled to show the characteristics of the data used when running analysis for each of the research questions. In the original data set, there were 43,726 student records. These records were filtered down to include data for only ninth-grade students who took Algebra I as a semester course for half of a credit (totaling one full credit per course per year), were identified as completing the course, and were not in the honors Algebra I courses. Additionally, any students in remedial or credit recovery courses were filtered out of the analysis for the general education courses.

For RQ1, when considering the district virtual versus the district f2f programs for the ninth-grade, first-time course takers who completed the course, data were filtered down to 26,730 records. From these records, I conducted a chi-squared test of independence to look for associations between course environment with end-of-course letter grades, ethnicity with end-of-course letter grades, and gender with end-of-course letter grades. Additionally, Cramer's V was calculated to understand the measure of an association, if one existed. For course environment and end-of-course letter grade, all expected cell frequencies were greater than five, and there was a statistically significant association, $\chi^2(3) = 37.684$, $p < .001$, Cramer's V = .038. For ethnicity and end-of-course letter grade, all expected cell frequencies were greater than five, and there was a statistically significant association, $\chi^2(12) = 3058.190$, $p < .001$. The association was small to moderate, Cramer's V = .195. For gender and end-of-course letter grade, all expected cell frequencies were greater than five, and there was a statistically significant association, $\chi^2(3) = 459.101$, $p < .001$. The association was small, Cramer's V = .131. Each of the statistically significant scores for the chi-squared test of independence indicated that grades were significantly associated with each of the variables, justifying the necessity of the independent variable and covariates in the data analysis.

Usually when running an ordinal logistic regression (OLR) calculation testing must first be conducted to ensure that the data does not fail certain assumptions which may skew the results in the OLR model. These assumptions could include those such as multicollinearity and proportional odds. Multicollinearity is necessary when there are two or more continuous independent variables. Due to the nature of this data being a census sample, and the independent variable and covariates being nominal or ordinal, assumption testing for multicollinearity was not needed. However, the assumption of proportional odds was tested.

In continuing with the ordinal logistics regression analysis, my next step was to conduct a full likelihood ratio test to determine if the assumption of proportional odds was met or violated. A full likelihood ratio test comparing the fit of the proportional odds location model to a model with varying location parameters was violated, $\chi^2(10) = 97.607$, $p < .001$, indicating the difference between the two models was large and statistically significant. To determine if the results of the full likelihood ratio test possibly flagged violations that did not exist, separate binomial regressions were conducted. Results indicated that the odds ratios for each run of the binomial regressions with cumulative dichotomous dependent variables were similar for each category of the independent variables indicating that there were flagged violations that did not exist.

Table 2 contains the results for the binomial regressions with cumulative splits of the dependent variable. Laerd Statistics (2015) stated that very large sample sizes can incorrectly flag assumption violations of the full likelihood ratio test. With the size of the sample for this study being over 26,000, this was most likely the case, and I made the decision to proceed with ordinal logistic regressions based on the similar outcomes of the separate binomial regressions.

Table 2

Binomial Regression for F2F Versus District Dependent Variable Cumulative Splits

		B	*SE*	χ^2	*df*	*Exp(B)*
		Cumulative Split F and A/B/C				
Step 1[a]	F2F	-0.776	0.152	26.192 ***	1	0.460
	Ethnicity			890.475 ***	3	
	White	-1.352	0.104	168.352 ***	1	0.259
	Hispanic	-0.307	0.049	38.462 ***	1	0.736
	Asian/ Multiracial	-1.756	0.063	772.389 ***	1	0.173
	Male	0.493	0.038	168.036 ***	1	1.637
	Constant	-0.969	0.153	40.208 ***	1	0.380
		Cumulative Split F/C and A/B				
Step 1[a]	F2F	-0.348	0.126	7.609 **	1	0.706
	Ethnicity			2341.319 ***	3	
	White	-1.323	0.054	594.739 ***	1	0.266
	Hispanic	-0.331	0.036	82.895 ***	1	0.719
	Asian/ Multiracial	-1.407	0.031	2058.790 ***	1	0.245
	Male	0.505	0.026	378.761 ***	1	1.656
	Constant	0.651	0.127	26.172 ***	1	1.917
		Cumulative Split F/C/B and A				
Step 1[a]	F2F	-1.121	0.221	25.632 ***	1	0.326
	Ethnicity			1389.023 ***	3	
	White	-1.404	0.059	566.065 ***	1	0.246
	Hispanic	-0.489	0.051	92.050 ***	1	0.613
	Asian/ Multiracial	-1.314	0.038	1193.936 ***	1	0.269
	Male	0.557	0.033	284.795	1	1.745
	Constant	2.997	0.223	180.939	1	20.033

Note. The greatest distance between odds ratios for f2f versus district virtual was 0.3798, White versus Black/African American was 0.0205, Hispanic versus Black or African American was 0.1227, Asian/Multiracial was 0.0959, and male versus female was 0.1078.

$^{**}p < .01$ $^{***}p < .001$

Although assumption testing for multicollinearity was not necessary for this research question due to population sampling, covariate patterns and cell frequencies were analyzed to ensure data represented a good model fit for the ordinal logistic regression analysis. Cell frequencies were analyzed for both data sets of the combination of Asian and Multiracial students and having these two ethnicity groups separated. When separated, of 80 covariate patterns 30% had frequencies under 5 (including the zero frequencies), with four cells (5%) having a value of 0. However, when Asian and Multiracial were combined into one category, there were 64 covariate patterns with 23.8% of them with frequencies under 5 (including the zero frequencies) and one cell (1.6%) having a value of 0. In both cases, the cases with zero to four cases came from the patterns for the district courses where there were far fewer students. The zero-frequency cell in the combined patterns represented female, Asian/Multiracial students taking the district courses and earning a grade of F. According to Laerd Statistics (2018), having no cells with zero frequency and 80% or more expected cell frequencies above 5, the overall goodness of fit measures can more reliably be interpreted.

Both the Pearson goodness-of-fit and the deviance goodness-of-fit tests indicated that the model was not a good fit for the data (Pearson $\chi^2(40) = 144.348$, $p < .001$; deviance $\chi^2(40) = 150.082$, $p < .001$. However, the final model statistically significantly predicted the dependent variable over and above the intercept-only model, $\chi^2(5) = 3542.680$, $p < .001$. As stated previously, due to census sampling and having all available data, I made the decision to continue with the procedure regardless of the higher percentage of covariate patterns with frequencies under 5 and poor Pearson and deviance model fit predictions. However, in doing so, results must be interpreted with caution. In addition to the final model, the omnibus tests of model effects indicated whether each of the independent variables had a statistically significant effect on the prediction of the grade earned. Gender had a statistically significant effect, Wald $\chi^2(1) = 517.394$, $p < .001$, course type had a statistically significant effect, Wald $\chi^2(1) = 29.73$, $p < .001$, and ethnicity had a statistically significant effect, Wald $\chi^2(3) = 2909.351$, $p < .001$. Based on these results, results of the odds ratios were compared to determine how the groups within the independent variables differed.

A cumulative odds ordinal logistic regression with proportional odds was run to determine the effect of course environment, gender, and ethnicity on end-of-course grades. The odds of a male earning higher grades was about half of that of a female, *Exp(B)* = .597, 95% CI [.571, .624], a statistically significant effect, Wald $\chi^2(1) = 517.394$, $p < .001$. The environment in which the course was taken presented higher odds of the f2f classes earning higher grades than the district virtual classes. The odds of a student earning a higher grade in the f2f courses was 1.815, 95% CI [1.465, 2.248] times that of a student in a district virtual class, a statistically significant effect, Wald $\chi^2(1) = 29.730$, $p < .001$.

As noted in Chapters 2 and 3, transactional distance is generally perceived as being higher in online courses than in f2f courses, often impairing the learning of online students. While the results found here must be weighed with caution due to the disproportionate sizes and demographics of the control and treatment groups, transactional distance can still be considered an aggravating factor in the poorer performance of the online students, especially with students in the f2f courses seeming to have had almost twice the odds in performing better than online students. Future analysis regarding the comparability of rigor and content between these environments and student perceptions could further solidify the extent to which transactional distance affects academic achievement.

Exploratory Analysis of RQ1 – F2F versus District Virtual

Regarding ethnicity, results indicated that the odds of White students earning a higher grade was 4.029, 95% CI [3.818, 4.252] times that of a Black/African American student, a statistically significant effect, Wald $\chi^2(1) = 2567.946$, $p < .001$. The odds of a Hispanic student earning a higher grade were about 50% better than a Black/African American student, Exp(B) = 1.465, 95% CI [1.347, 1.531], and was a statistically significant effect, Wald $\chi^2(1) = 122.326$, $p < .001$. The odds of a student in the Asian/Multiracial ethnic category earning a higher grade were 3.950, 95% CI [3.602, 4.330] times that of a Black/African American student, a statistically significant effect, Wald $\chi^2(1) = 855.306$, $p < .001$. (See Table 3 for the parameter estimates for this data.)

Table 3

Ordinal Logistic Regression Parameter Estimates F2F Versus District Virtual Course Grades

Variable		*B*	*SE*	95% CI for *B*	$\chi^{2\,a}$	*Exp(B)*	95% CI for *Exp(B)*
Threshold	F	-1.188	.1106	[-1.405, -.971]	115.433	.305 ***	[.245, .379]
	C	.902	.1103	[2.292, 2.728]	66.856	2.464 ***	[1.985, 3.059]
	B	2.510	.1113	[2.292, 2.728]	508.740	12.307 [b]	[9.895, 15.306]
F2F [c]		0.596	0.1093	[0.382, 0.810]	29.728	1.815 ***	[1.465, 2.248]
Asian/Multi [d]		1.374	0.0470	[1.282, 1.466]	855.306	3.950 ***	[3.602, 4.330]
Hispanic [d]		.362	.0327	[.298, .426]	122.326	1.436 [e]	[1.347, 1.531]
White [d]		1.394	.0275	[1.340, 1.447]	2567.946	4.029 ***	[3.818, 4.252]
Male [f]		-.516	0.0227	[-.561, -.472]	517.394	.597 ***	[.571, .624]
(Scale)		1 [g]					

Note. CI = Confidence Interval. Dependent variable was Letter Grade. The model analyzed the relationship between the type of course, ethnicity, gender, and letter grade.

[a] df = 1. [b]*p* = 0.704. [c]*p* = 0.680. ****p* < .001.

[d] Reference category was District Virtual. [e] Reference category was Black/African American. [f] Reference category was Female. [g] Fixed at the displayed value.

Interesting results appeared when additional ordinal logistic analyses were conducted with ethnicity being recoded into several dichotomous variables to compare each group individually to a sum of all other groups, for example, White to all others. Results from these calculations indicated that the odds of earning a higher end of course grades were statistically significantly better for non-Black/African Americans than a Black/African American, non-Hispanic than Hispanic, White than non-White, and Asian/Multiracial than non-Asian/Multiracial. Specifically, the odds of a non-Black/African American student (Table 4) earning a higher overall course grade was

2.902, 95% CI [2.772, 3.037] times higher than a Black/African American student earning a higher grade, Wald $\chi^2(1) = 2103.665$, $p < .001$. The odds of a non-White student (Table 5) earning a higher overall course grade was .312, 95% CI [.296, .328] times that of a White student earning a higher grade, Wald $\chi^2(1) = 2085.724$, $p < .001$. The odds of a non-Hispanic student (Table 6) earning a higher overall course grade was 1.243, 95% CI [1.171, 1.320] times that of a Hispanic student, Wald $\chi^2(1) = 50.357$, $p < .001$. Finally, the odds of a non-Asian/Multiracial student (Table 7) earning a higher overall course grade was .438, 95% CI [.401, .479] times higher than an Asian/Multiracial student, Wald $\chi^2(1) = 335.828$, $p < .001$.

Table 4

Ordinal Logistic Regression Parameter Estimates Table for Non-Black/African American Versus Black/African American

Variable		*B*	*SE*	95% CI for *B*	$\chi^{2\,a}$	*Exp(B)*	95% CI for *Exp(B)*
Threshold	F	-1.228	.1103	[-1.444, -1.012]	124.063	.293 ***	[.236, .363]
	C	.829	.1100	[.614, 1.045]	56.874	2.292 ***	[1.847, 2.843]
	B	2.392	.1108	[2.175, 2.609]	465.757	10.935 ***	[8.800, 13.588]
F2F [b]		.520	.1089	[.306, .733]	22.774	1.682 ***	[1.358, 2.082]
Not Black/ African American [c]		1.065	.0232	[1.020, 1.111]	2103.665	2.902 ***	[2.772, 3.037]
Male [d]		-.501	.0226	[-.545, -.457]	491.196	.606 ***	[.580, .633]
(Scale)		1[b]					

Note. CI = Confidence Interval. Dependent variable was Letter Grade. The model analyzed the relationship between the type of course, ethnicity, gender, and letter grade.

[a] df = 1

[b] Reference category was District Virtual. [c] Reference category was Black/African American. [d] Reference category was Female.

[e] Fixed at the displayed value.

*** $p < .001$

Table 5

Ordinal Logistic Regression Parameter Estimates Table for Non-White Versus White

Variable		*B*	*SE*	95% CI for *B*	$\chi^{2\,a}$	*Exp(B)*	95% CI for *Exp(B)*
Threshold	F	-2.532	.1114	[-2.750, -2.313]	516.800	.080 ***	[.064, .099]
	C	-.485	.1104	[-.702, -.269]	19.356	.615 [b]	[.496, .764]
	B	1.085	.1105	[.869, 1.302]	96.555	2.961 ***	[2.384, 3.677]
F2F [c]		.576	.1091	[.363, .790]	27.922	1.779 ***	[1.437, 2.203]
Not White [d]		-1.166	.0255	[-1.216, -1.116]	2085.724	.312 ***	[.296, .328]
Female [e]		0.478	0.0224	[0.434, 0.522]	455.127	1.613 ***	[1.544, 1.686]
(Scale)		1 [f]					

Note. CI = Confidence Interval. Dependent variable was Letter Grade. The model analyzed the relationship between the type of course, ethnicity, gender, and letter grade.

[a] df = 1. [b] $p = 0.178$

[c] Reference category was District Virtual. [d] Reference category was White. [e] Reference category was Female.

[f] Fixed at the displayed value.

*** $p < .001$

Table 6

Ordinal Logistic Regression Parameter Estimates Table for Non-Hispanic Versus Hispanic

Variable		*B*	*SE*	95% CI for *B*	$\chi^{2\,a}$	*Exp(B)*	95% CI for *Exp(B)*
Threshold	F	-1.567	.1123	[-1.787, -1.347]	194.808	.209 ***	[.167, .260]
	C	.387	.1117	[.168, .606]	12.026	1.473 ***	[1.183, 1.834]
	B	1.865	.1123	[1.645, 2.086]	275.922	6.459 ***	[5.183, 8.049]
F2F [b]		.429	.1083	[.217, .641]	15.677	1.536 ***	[1.242, 1.899]
Not Hispanic [c]		.218	.0307	[.158, .278]	50.357	1.243 ***	[1.171, 1.320]
Male [d]		-.478	.0224	[-.522, -.434]	455.127	.620 ***	[.593, .648]
(Scale)		1 [e]					

Note. CI = Confidence Interval. Dependent variable was Letter Grade. The model analyzed the relationship between the type of course, ethnicity, gender, and letter grade.

[a] df = 1.

[b] Reference category was District Virtual. [c] Reference category was Hispanic. [d] Reference category was Female.

[e] Fixed at the displayed value.

*** $p < .001$

Table 7

Ordinal Logistic Regression Parameter Estimates Table for Non-Asian/Multiracial Versus Asian/Multiracial

Variable		*B*	*SE*	95% CI for *B*	$\chi^{2\,a}$	Exp(B)	95% CI for *Exp(B)*
Threshold	F	-2.541	.1174	[-2.771, -2.311]	468.488	.079 ***	[.063, .099]
	C	-.577	.1165	[-.805, -.349]	24.548	.562 ***	[.447, .706]
	B	.914	.1165	[.686, 1.142]	61.553	2.949 ***	[1.985, 3.134]
F2F [b]		.424	.1084	[.212, .636]	15.307	1.528 ***	[1.236, 1.890]
Not Asian/ Multiracial [c]		-.825	.0450	[-.913, -.737]	335.828	.438 ***	[.401, .479]
Male [d]		-.487	.0225	[-.532, -.443]	471.145	.614 ***	[.588, .642]
(Scale)		1[e]					

Note. CI = Confidence Interval. Dependent variable was Letter Grade. The model analyzed the relationship between the type of course, ethnicity, gender, and letter grade.

[a] df = 1.

[b] Reference category was District Virtual. [c] Reference category was Asian/Multiracial.

[d] Reference category was Female.

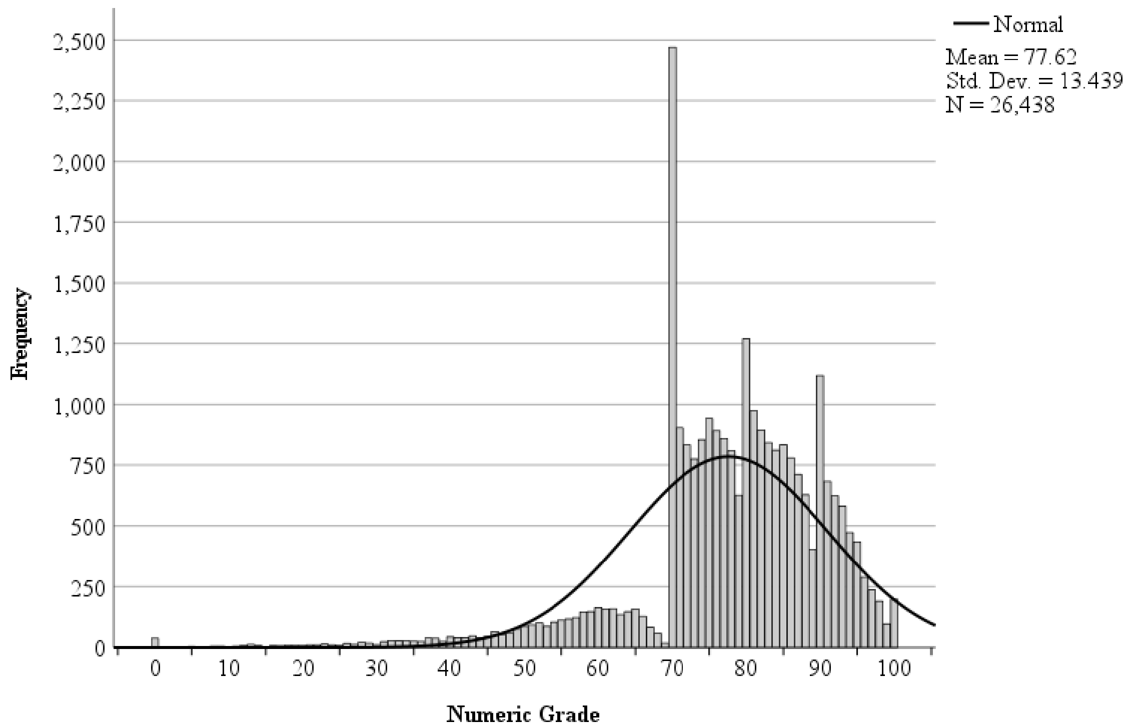
[e] Fixed at the displayed value.

*** $p < .001$

Deeper consideration of the f2f and district virtual courses was conducted to get a further understanding of the differences found in the previous calculations. Of the 26,730 records, 98.9% took Algebra I in a f2f course, while only 1.1% took the course virtually. Distribution statistics for f2f students based on actual grades earned (0 to 100) resulted in the skewness of -1.583 indicating a negatively skewed distribution, and a positive kurtosis of 4.711 indicating a lightly-tailed distribution. Distribution statistics for district virtual students based on actual grades earned (0 to 100) resulted in the skewness of -1.782 indicating a negatively skewed distribution, and a roughly normal kurtosis of 2.095. Both distributions indicated that more students were passing than were not passing, a pleasant finding in educational research. However, in both environments, there was a spike in grades at the grade of exactly 70% (n = 2480), compared to a total of 1,217 for grades 60-69% (n = 20 for 69%) and 914 students for 71%. Students who scored a 70% or higher were considered to have passed the class while grades of 69% and below were failing. Figure 2 shows the histogram of grade distributions for the f2f courses while Figure 3 shows a histogram of the grades earned in the district virtual courses.

Figure 2

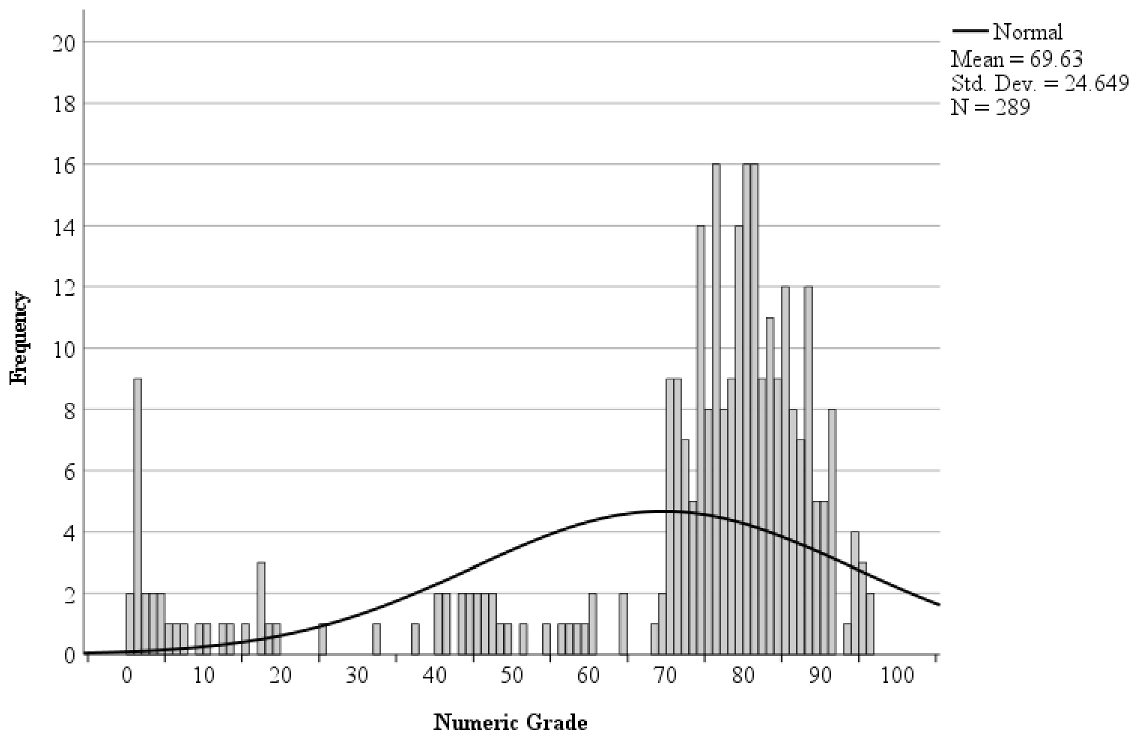
Grade Distribution for Ninth-Grade Algebra I F2F Courses During the School Years 2016-17 to 2019-20.



Note. The largest spike in the frequencies was at 70% (*n* = 2471). The next two major spikes were at 80% (*n* = 1270) and 90% (*n* = 1119). In scores below 70%, there was a minor spike at 0% (*n* = 38). Scores steadily increase in average frequency up to 60% (*n* = 164), where average frequencies drop until 70%.

Figure 3

Grade Distribution for Ninth-Grade Algebra I District Virtual Courses During the School Years 2016-17 to 2019-20.

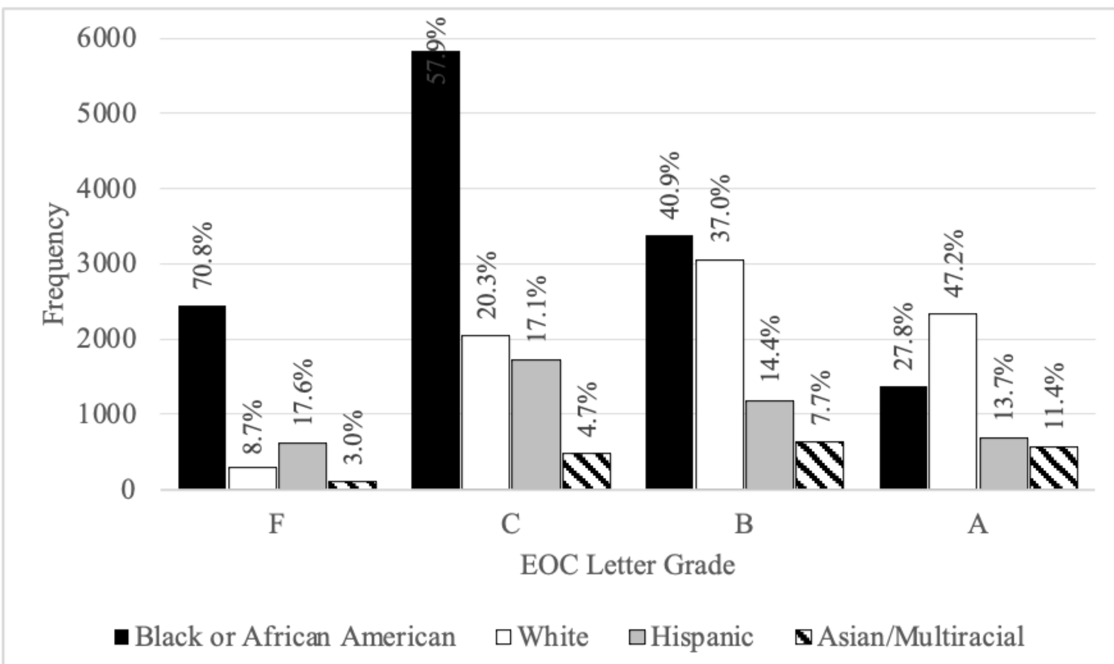


Note. Scores begin to be consistently above the distribution curve at 70% (*n* = 9) and 71% (*n* = 9). The most frequent scores, all with counts equal to 16, were 76%, 80%, and 81%. In scores lower than 70%, the scores with the highest frequency were 1% (*n* = 9) and 17% (*n* = 3).

Post hoc analyses were conducted to determine the distribution of grades among the different ethnicity groupings and genders. Crosstabulation between numeric grade and ethnicity groups was conducted to reveal which ethnic group accounted for the highest percentages of each grade. Black/African American students accounted for over half (65.4%) of the end-of-course grades of 70% with White students accounting for 12.7%, Hispanic students for 18.5%, and Asian/Multiracial students for 3.3%. Males accounted for slightly more than half (56.5%) of the course grade of 70%. Figure 4 and Figure 5 show the distribution of grades in terms of final letter grades by ethnicity, and the distribution of grades in terms of final letter grades by gender, respectively.

Figure 4

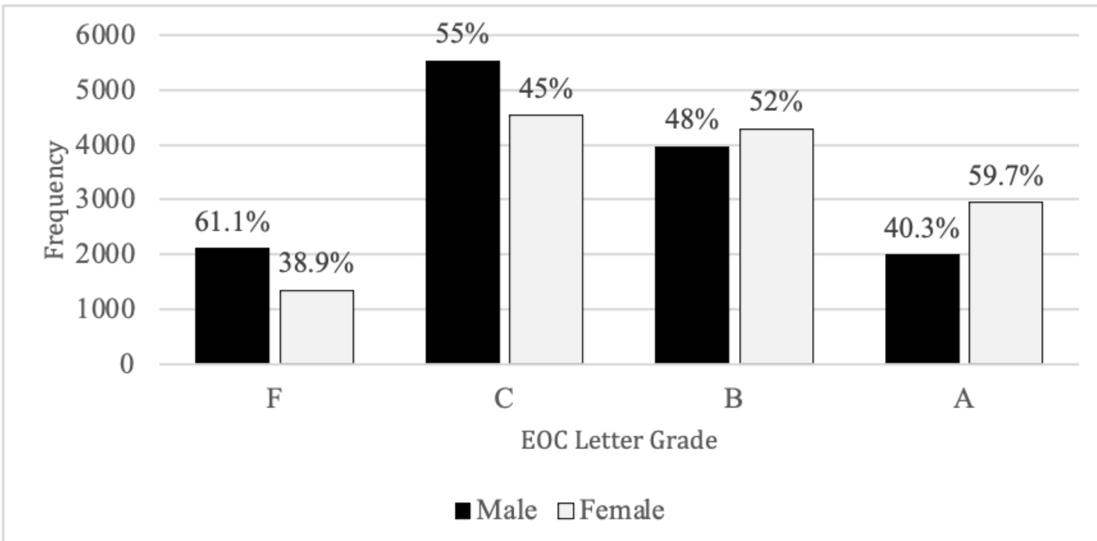
Letter Grade Distribution for Ethnicity Groups in F2F and District Virtual Courses



Note. Percentages indicated the proportion that each ethnicity accounted for each letter grade category. Black/African American students accounted for much higher percentages of Cs and Fs (61.2%) than As and Bs (36%), while White students accounted for a higher percentage of As and Bs (40.8%) than Cs and Fs (17.3%).

Figure 5

Letter Grade Distribution for Gender Groups in F2F and District Virtual Courses



Note. Percentages indicated the proportion each gender accounted for each letter grade category. Male students accounted for a slightly higher percentage of the Cs and Fs (56.5%) and a slightly lower percentage of As and Bs (45.1%), while Females accounted for 43.5% of Cs and Fs and 54.9% of As and Bs.

Imitating results of previous studies, the findings of this study for the comparisons of a f2f Algebra I course being taken by ninth-grade students for the first time versus the same courses in the district's virtual course environment suggested that Black/African American students were disproportionately performing worse than students of other ethnicities. Black/African American students accounted for 70.8% of all Fs in the district, and 18.8% of all Black/African American students failed Algebra I. Additionally, male students accounted for 61.1% of all Fs in the district, and 15.5% of all males failed Algebra I. (See Table 8 for a breakdown of percentages of grades for each ethnicity and gender.) While White students made up only 28.9% of the population of all ninth-grade students taking Algebra I, they accounted for 40.8% of the higher grades, A and B. Between the two environments specifically, students in the f2f courses had higher odds of receiving higher grades. Over the school years of 2016-17 to 2019-20, students in the f2f courses had an overall average of 77.62% while students in the district virtual courses had an overall average of 69.63%, a difference of about 8%. The result of students in the virtual courses performing lower than students in the f2f courses was in line with the results of past research (Ahn & McEachin, 2017; Hart et al., 2019; Heinrich et al., 2019; Heppen et al., 2017; Hughes et al., 2015). Implications of these results will be discussed in further detail in Chapter 5.

Table 8

Percentage of Students with EOC Letter Grades by Ethnicity and Gender Grouping

		Ethnicity/ Gender Group (%)			
Comparison Group	Grade	Black/ African American	White	Hispanic	Asian/ Multiracial
		Male			
Specific	A	8.1	24.5	11.1	28.6
All		10.8	19.2	4.7	5.6
Specific	B	22.7	41.0	25.3	34.5
All		18.3	19.3	6.4	4.1
Specific	C	47.0	30.0	45.3	30.2
All		31.1	11.6	9.5	2.9
Specific	F	22.3	4.6	18.3	6.7
All		18.3	5.1	11.1	1.9
		Female			
Specific	A	13.2	36.0	21.2	35.3
All		17.0	28.0	9.0	5.8
Specific	B	29.3	38.1	31.3	37.5
All		22.6	17.8	7.9	3.7
Specific	C	42.4	22.8	36.9	22.5
All		26.9	8.7	7.6	1.8
Specific	F	15.1	3.2	10.7	4.7
All		27.8	3.6	6.5	1.1

Note. "Specific" indicates the percentage of that specific group who earned that grade. For example, 8.1% under Black/African American males for grade A indicates that 8.1% of Black/African American males earned an A. "All" indicates the percentage that specific group earned for all grades of that category. For example, 10.8% under Black/African American males indicates that Black/African American males earned 10.8% of all As earned during the school years 2016-17 to 2019-20.

Overall findings of ordinal logistic regressions with cumulative odds indicated that students who were female, White, or Asia/Multiracial had higher odds of receiving better grades than students who were male, Black/African American, or Hispanic. Additionally, a student taking an Algebra I course in the f2f environment has higher odds of receiving better grades than students taking the course virtually. These findings mimic those of previous research studies (Ahn & McEachin, 2017; Hart et al., 2019; Heinrich et al., 2019; Heppen et al., 2017; Hughes et al., 2015). This will be discussed in further detail in Chapter 5.

Analysis for RQ1 - District versus State Virtual

In the original data set, there were 43,726 student records. Again, these records were filtered down to include data for only ninth-grade students who took Algebra I as a semester course for half of a credit (totaling one full credit per course per year), were not in the honors Algebra I courses, and were not in remedial or credit recovery courses. For this portion of research question one, only students who took the courses virtually, either in the district virtual course or the state virtual course, were considered. The filtered total for these cases was 306, 94.4% of which were from the district courses and 5.6% from the state courses. As noted earlier, the state virtual courses were not proportional to that of the district as a whole for ninth-grade Algebra I students as there were only 17 students, 15 (88.8%) of which were White students (1 was Black or African American and 1 was Hispanic). Of the 17 students, only 5 (29.4%) were female. At the lower end of the recommended EPV for logistic regression, 10 EPV, there would need to be at least 30 students to perform a reliable analysis. Several researchers have suggested that regression calculation problems are minimized when EPV is ten or greater (Grant et al., 2019; Halinski & Feldt, 1970; Miller & Kunce, 1973; Peduzzi et al., 1996). With the data available for the state virtual courses being maximized at about five EPV, it was not feasible to conduct accurate calculations for any group comparisons. Additionally, analyzed cell frequencies indicated that of the 64 covariate patterns, 38 cells (59.4%) had fewer than five cases with 21 cells having zero frequencies (32.8% of total). This was not surprising considering the small number of overall data points. Instead of proceeding with regression analysis, a description of the data was provided. For RQ2 and RQ3, regression analysis will continue to be utilized for the district virtual and f2f courses, but not for the state virtual courses.

For the state virtual courses, of the 17 students who completed the course, four (23.5%) of them failed, four (23.5%) passed with a C, seven (41.2%) passed with a B, and two (11.8%) passed with an A. The overall numeric grade average was 71.88%, and the highest grade earned was 93%. The one Black/African American student was a male who passed with a C (72%), while the Hispanic male failed with an overall grade of 12%. Of the five White females, one female failed with a 60%, one female earned an A (93%), two passed with a B, and one passed with a C. Of the 10 White males, one passed with an A, five passed with a B, two passed with a C, and two failed.

A cumulative odds ordinal logistic regression with proportional odds was run to determine the effect of course environment, gender, and ethnicity on end-of-course grades when comparing f2f to all virtual students by combining the district and state data into one category. Results differed only slightly, less than 0.03, from the previous calculations comparing f2f to district virtual only for all comparison groups. Unfortunately, there were not enough state virtual data for this district to construct any meaningful analysis or inferences.

Analysis for RQ2 – F2F versus District Virtual

The goal of RQ2 was to identify if there was a difference in student FastBridge assessment scores as determined by their ROI and raw scores. As previously discussed, FastBridge data were only included for one school year. Of the original 43,726 records, 1061 met the criteria of ninth-grade students taking Algebra I for the first time and had raw scores recorded for their spring assessment. F2F students' scores accounted for 99.2% of this data while there were only nine cases for virtual students, none of which were from the state district data. Due to the lack of data for virtual students when comparing F2F to the district virtual courses, regression analysis could not be conducted to compare the two environments. Instead, exploratory analysis using ordinal logistic regression was conducted to calculate odds ratios for gender and ethnicity within the f2f environment only.

Of the nine district virtual students with FastBridge data reported, there were three females (two Black/African American and one Hispanic), and six males (three Black/African American, two White, and one Asian/Multiracial). While growth scores for the entire data set of 1,061 students range from -3.16 to 1.65, students in the district virtual course ranged from -0.83 to 0.09, five of which were negative implying learning loss. Two ROI scores were positive indicating learning gained but were both less than 0.1 indicating an extremely small amount of learning gained. Two of the ROI scores were null, indicating these students did not take a fall assessment. Additionally, the average spring raw score for all included f2f and virtual district students was 227.7517 while the average for the district virtual group only was 224.8278. See Table 9 for further descriptive information for the district virtual courses FastBridge assessment data.

Table 9

District Virtual FastBridge ROI and Spring Raw Scores by Ethnicity and Gender

Ethnicity	Gender	EOC Grade (%)	ROI	Spring Raw Score
Black/African American	Male	76 [a]	-1.01	216.25
Black/African American	Male	82	-0.04	216.26
Black/African American	Male	88 [a]	NULL	222.06
White	Male	85 [a]	0.09	227.21
White	Male	88	-0.47	210.15
Asian/Multiracial	Male	86	NULL	232.64
Black/African American	Female	19	0.04	235.82
Black/African American	Female	48	-0.83	232.24
Hispanic	Female	60	-0.17	230.78

Note. EOC grades were included as a reference when reviewing ROI and spring raw scores. Data was sorted by gender and ethnicity.

[a] Students were flagged as “Instructional Setting Special Ed”. These students were included to keep the data as extensive as possible.

Due to a lack of data in both the district and state virtual data, ordinal logistic regressions could not be conducted to compare the online programs to the f2f courses, and TDT could not be used as part of the discussions. There was no way to know whether the course environment played a part in students' standardized assessment success.

Exploratory Analysis RQ2 – Gender and Ethnicity Differences Within the F2F Environment

For the f2f FastBridge assessment data, two separate cumulative odds ordinal logistic regressions with proportional odds were conducted. The first utilized ROI growth data as the dependent variable while the second analysis was conducted using the spring raw scores. Since both the ROI and raw scores were naturally continuous variables, they were both broken into ordinal groups based on their means and standard deviations. For example, ROI data were binned into four groups based on $M = .2038$, $SD = .53847$. These groups represented moderate to large learning loss ($\leq$ -0.33, $n = 102$), small learning loss to small learning gain (-0.32 to 0.2, $n = 342$), moderate learning gain (0.21 to 0.74, $n = 349$), and large learning gains (> 0.74, $n = 112$).

A full likelihood ratio test was conducted. Results indicated the test was violated, $\chi^2(2) = 17.788$, $p < .001$, meaning that the difference between the two models was large and statistically significant. To determine if the results of the full likelihood ratio test possibly flagged violations that did not exist, separate binomial regressions were conducted. The odds ratio for Hispanic students was similar across each dichotomous grouping of the dependent variable, however, the odds ratios were substantially different for each of the other ethnicities and gender. Instead of continuing with ordinal logistic regression, multinomial logistic regression was utilized (Laerd Statistics, 2015).

To conduct a multinomial logistic regression, as with the ordinal logistic regressions, certain assumptions should be met. The first four assumptions regarding having nominal dependent variables, having one or more independent variables at any level, independence of observations, and having no multicollinearity had already been established. According to Laerd Statistics, the next two assumptions should show that "there needs to be a linear relationship between any continuous independent variables and the logit transformation of the dependent variable…[and] there should be no outliers, high leverage values or highly influential points" (2018, *Assumptions*). In this data set, there were no continuous independent variables so there was no need to conduct testing for linear relationships between continuous variables and logit transformations of the dependent variable. After grouping the dependent variable into the four categories, no ROIs could become outliers, leverage points, or highly influential points.

Both the Pearson goodness-of-fit and the deviance goodness-of-fit tests indicated that the model was a good fit for the data (Pearson $\chi^2(9) = 6.944$, $p = .643$; deviance $\chi^2(9) = 6.860$, $p = .652$. Additionally, the final model statistically significantly predicted the dependent variable over and above the intercept-only model, $\chi^2(12) = 54.3$, $p < .001$. The likelihood ratio test of this data for a multinominal regression analysis indicated that both ethnicity, $\chi^2(9) = 33.267$, $p < .001$, and gender, $\chi^2(3) = 23.342$, $p < .001$, statistically significantly predicted the ROI category. Results of the odds ratios were compared to determine how the groups within the independent variables differed (see Table 10 for more information).

Table 10

Multinominal Logistic Regression Parameter Estimates F2F ROI Scores by Ethnicity and Gender

Dependent Variable [a]	Independent Parameters	B	SE	$\chi^{2\,b}$	*p*	Exp(B)	95% CI for Exp(B)
Moderate to Large Loss	Intercept	0.204	0.264	0.598	0.439		
	White [c]	-0.351	0.318	1.217	0.270	0.704	[0.377, 1.313]
	Hispanic [c]	-0.021	0.492	0.002	0.965	0.979	[0.373, 2.568]
	Asian/ Multiracial [c]	-1.120	0.483	5.364	0.021	0.326	[0.127, 0.842]
	Male [d]	-0.001	0.281	0.000	0.996	0.999	[0.576, 1.732]
Small Loss to Small Gain [c]	Intercept	0.802	0.228	12.405	0.000		
	White [c]	-0.388	0.267	2.118	0.146	0.678	[0.402, 1.144]
	Hispanic [c]	0.682	0.400	2.902	0.088	1.978	[0.902, 4.335]
	Asian/ Multiracial [c]	-0.213	0.335	0.405	0.525	0.808	[0.419, 1.558]
	Male [d]	0.835	0.224	13.919	0.000	2.304	[1.486, 3.572]
Moderate Gain [c]	Intercept	0.984	0.225	19.148	0.000		
	White [c]	0.077	0.262	0.087	0.767	1.080	[0.647, 1.804]
	Hispanic [c]	0.503	0.404	1.549	0.213	1.653	[0.749, 3.650]
	Asian/ Multiracial [c]	-0.776	0.357	4.721	0.030	0.460	[0.228, 0.927]
	Male [d]	0.351	0.222	2.505	0.114	1.421	[0.920, 2.194]

Note. CI = Confidence Interval.

[a] Reference category was Large Gain.

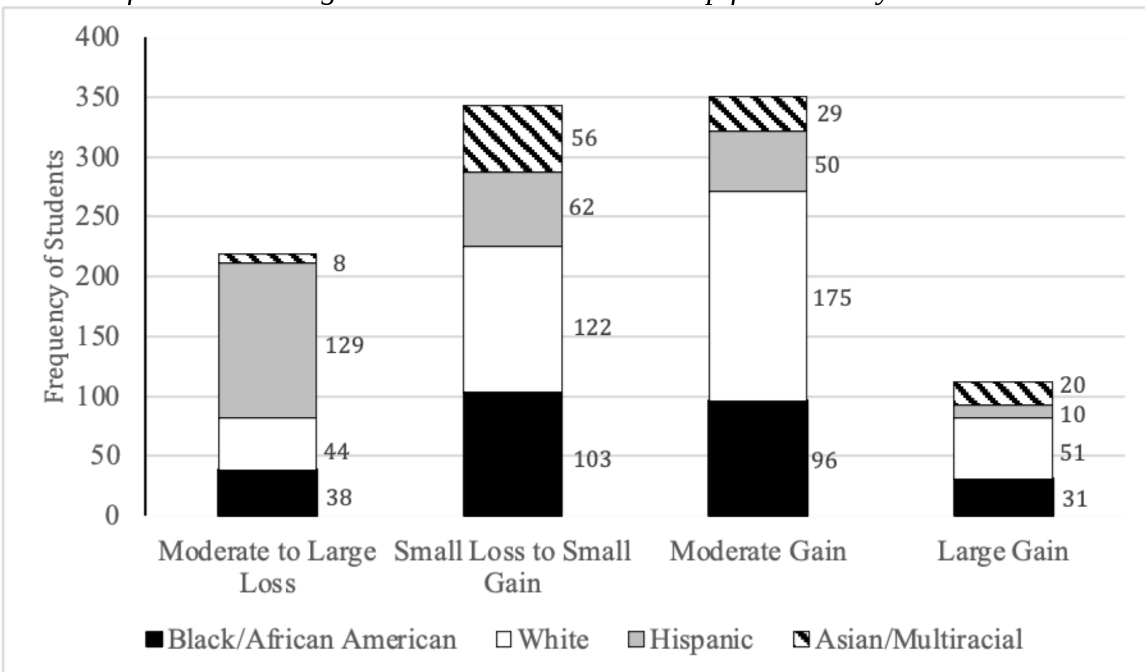
[b] df = 1

[c] Reference category was Black/African American. [d] Reference category was Female.

Statistically significant effects were calculated when comparing the odds of Asian/Multiracial students to Black/African American students for both the moderate to large learning loss and moderate learning gain categories. Both statistically significant effects indicated that the odds of Asian/Multiracial students scoring lower as opposed to higher was less than half that of a Black/African American student. The odds of Asian/Multiracial students having moderate to large learning loss rather than high learning gains was .326, 95% CI [0.127, 0.842], Wald $\chi^2(1) = 5.364$, $p < .05$, of Black/African American students. The odds of Asian/Multiracial students scoring in the moderate learning gain category versus the large gains category was .460, [0.228, 0.927], times that of Black/African American students, also a statistically significant effect, Wald $\chi^2(1) = 4.721$, $p < .05$. Finally, male students were over twice as likely to be in the small loss to small gain category, as opposed to the large learning gain category, than female students, *Exp(B)*= 2.304, [1.486, 3.572], a statistically significant effect, Wald $\chi^2(1) = 13.919$, $p < .001$. No other statistically significant effects were found in ROI. See Figure 6 and Figure 7 for ROI distributions amongst ethnicities and genders, respectively.

Figure 6

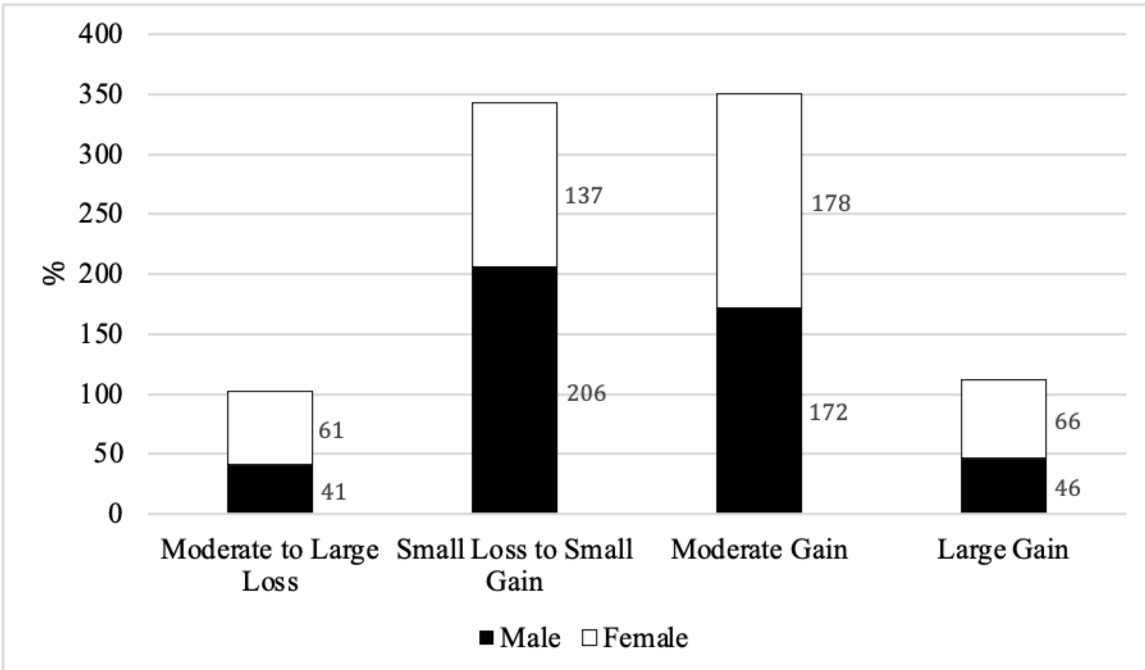
Number of F2F FastBridge Distribution Per ROI Group for Ethnicity



Note. Black/African American n = 268, White n = 392, Hispanic n = 251, Asian/Multiracial n = 113.

Figure 7

Number of F2F FastBridge Distribution per ROI Group for Genders



Note. Male *n* = 465, Female *n* = 318.

For the f2f spring raw scores, a total of 1,052 first-time Algebra I ninth-grade students had accompanying scores. Ordinal groupings were originally considered based on Illuminate Education's *FastBridge aMath Score Interpretation Guide* (2019). This guide provides a breakdown of students' expected skills based on a score range and recommends strategies for future instruction based on the range categories. "The aMath scores are listed according to bands of about 50 points each starting at 145 and ending at 275" (Illuminate Education, 2019, p. 2). The guide has categories of 145-200, 200-250, and 250-275. However, using this guide for ordinal categories would put 37 students in the first category, 1,014 in the second category, and zero students in the third category. Instead of splitting the data according to the interpretation guide, data was split in the same manner as the ROI scores with M = 227.78 and SD =12.39. With this strategy, four categories were created that roughly represent a normal distribution curve (skewness = -.260, SE = .075, kurtosis = -.666, SE = .151). The categories were Low ($\leq$ 215.39, n = 152), Lower Middle (215.4 to 227.78, n = 294), Upper Middle (227.79 to 240.17, n = 477), and High (> 240.18, n = 129).

As with the previous analysis, there was no need to test for multicollinearity due to having zero continuous independent variables. A full-likelihood ratio test indicated that the assumption of proportional odds was violated, $\chi^2(8) = 26.083$, $p < .001$. Separate binomial regressions with cumulative splits on the dependent variable indicated that the full-likelihood ratio test may have been violated. For each of the independent variables, the largest of the odds ratios was approximately twice that of the odds ratio for the equivalent factor, though all odds were less than 1. For example, the greatest difference between odds ratios exists for males to females for the binomial regressions of the first and third cumulative splits (0.962 to 0.451). As a precaution, multinomial logistic regression was used to ensure results were more reliable. (See Table 11 for results of the separate binomial logistic regression for this data).

Table 11

Binomial Regression for Cumulative Splits on the DV for F2F Spring Raw Scores

		B	*SE*	χ^2	*df*	*Exp(B)*
		Cumulative Split at Low and LM/UM/High				
Step 1	Ethnicity			88.312	3	
	White [a]	-2.597	0.304	72.842	1	0.074 ***
	Hispanic [a]	-0.905	0.259	12.265	1	0.404 ***
	Asian/ Multiracial [a]	-1.660	0.365	20.666	1	0.190 ***
	Male [b]	-0.039	0.187	0.043	1	0.962 [c]
	Constant	-0.831	0.150	30.522	1	0.436 ***
		Cumulative Split at Low/LM and UM/High				
Step 1	Ethnicity			158.584	3	
	White [a]	-1.918	0.161	142.706	1	0.147 ***
	Hispanic [a]	-1.160	0.202	32.874	1	0.313 ***
	Asian/ Multiracial [a]	-1.853	0.238	60.683	1	0.157 ***
	Male [b]	-0.219	0.137	2.541	1	0.804 [d]
	Constant	0.916	0.136	45.204	1	2.498 ***
		Cumulative Split at Low/LM/UM and High				
Step 1	Ethnicity			59.265	3	
	White [a]	*-2.017*	0.332	36.999	1	0.133 ***
	Hispanic [a]	-0.358	0.519	0.477	1	0.699 [e]
	Asian/ Multiracial [a]	-2.462	0.372	43.886	1	0.085 ***
	Male [b]	-0.797	0.209	14.619	1	0.451 ***
	Constant	3.950	0.340	135.124	1	51.910 ***

Note. LM = Lower Middle, UM = Upper Middle.

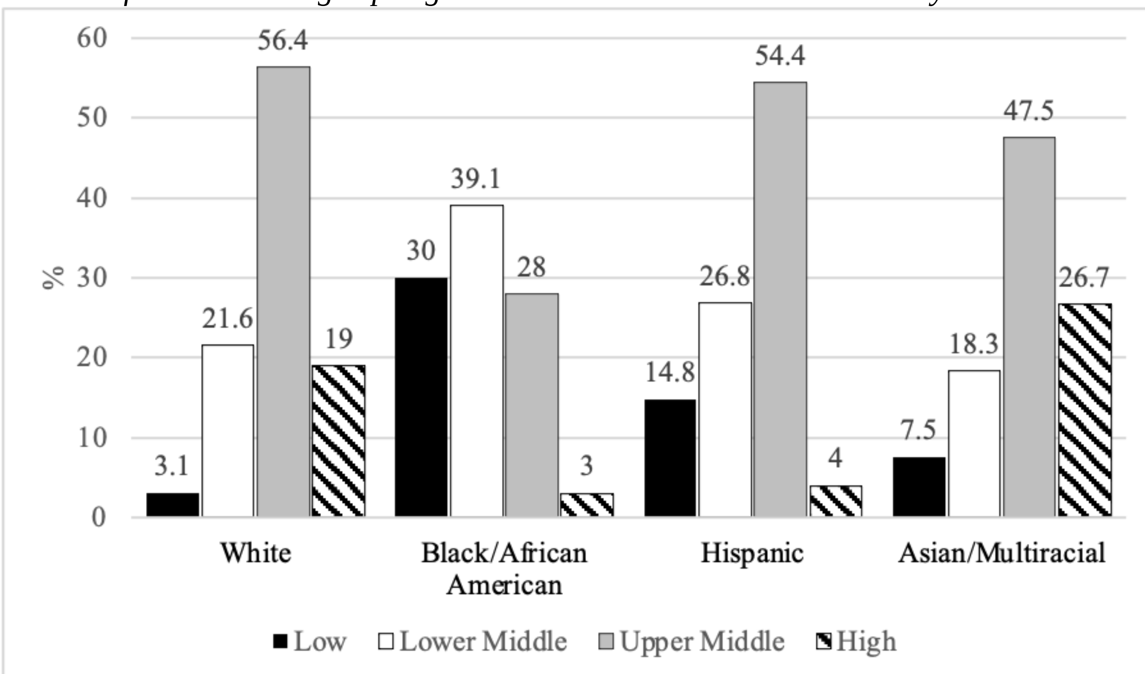
[a] Reference category was Black/African American. [b] Reference category was Female.

[c] *p* = .835. [d] *p* = .111. [f] *p* = .490.

Based on the dataset, all assumptions of multinomial regressions were either already met or did not need to be tested for. According to the goodness-of-fit test, the data for the spring raw Theta scores, after being converted to an ordinal variable, did not fit the model (Pearson $\chi^2(9) = 37.268$, $p = .001$ and deviance $\chi^2(9) = 39.274$, $p < .001$. However, overall model fit indicated that the model statistically significantly predicts the dependent variable better than the intercept-only model, $\chi^2(12) = 264.093$, $p < .001$, and likelihood ratio tests suggested that both ethnicity ($\chi^2(9) = 246.461$, $p < .001$) and gender ($\chi^2(3) = 16.015$, $p < .001$) statistically significantly affected the dependent variable. Laerd Statistics (2018) suggests that either the Pearson and deviance goodness-of-fit tests or the overall model fit can be used to determine if the data fits the model. Based on the overall model fit and likelihood ratio tests, I made the decision to proceed with multinomial logistic regression analysis. See Figure 8 and Figure 9 for the percent of f2f FastBridge spring Theta score distributions per ethnicity and gender, respectively.

Figure 8

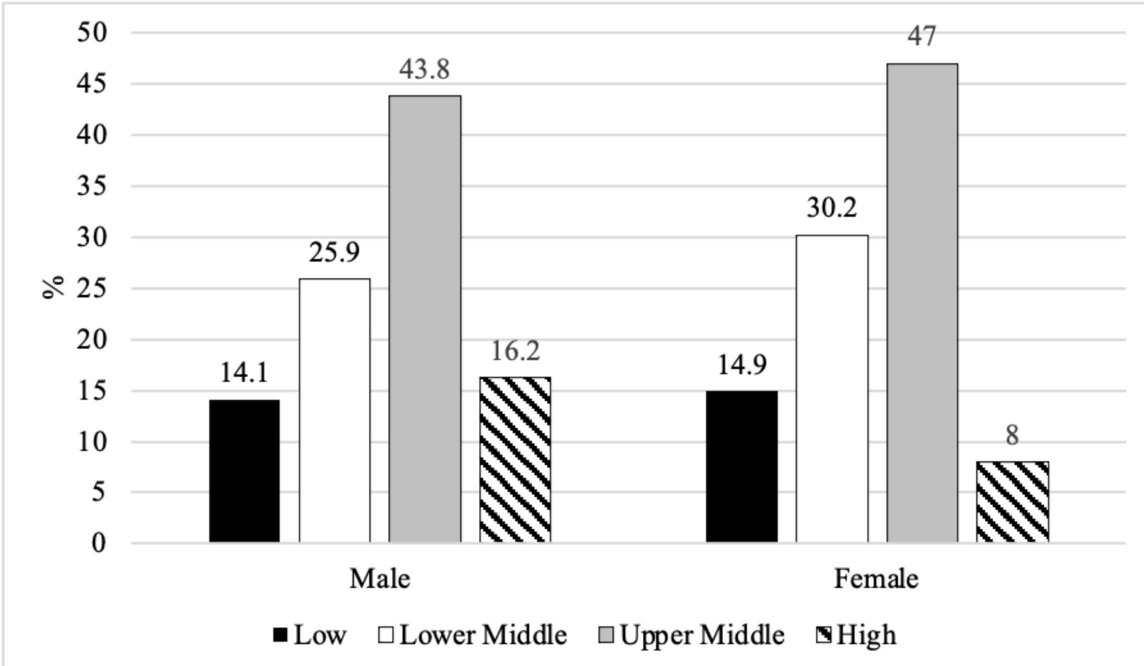
Percent of F2F FastBridge Spring Theta Score Distribution Per Ethnicity



Note. Percentages indicated the proportion of each spring raw score category for each ethnicity. All percentages were rounded to the nearest tenth.

Figure 9

Percent of F2F FastBridge Spring Theta Score Distribution Per Gender



Note. Percentages indicated the proportion of each spring raw score category for each gender. All percentages were rounded to the nearest tenth.

Multiple multinomial logistic regressions were conducted to make comparisons between each grouping of pairs of the dependent variable on each matched group of ethnicities (see Table 12 for a full multinomial regression table). For example, an analysis was conducted comparing each of the dependent categories in comparison to scoring Low while holding Black/African American students as the reference category for ethnicity, then again for White students as the reference, then Hispanic, and finally

Asian/Multiracial. Analysis was repeated to compare each of the dependent categories to Lower Middle, then Upper Middle, and finally, to High.

Table 12

Multinomial Logistic Regressions for F2F Spring Raw Scores

Spring Theta Categories		*B*	*SE*	χ^2	*p*	*Exp(B)*	95% CI for *Exp(B)*
Lower Middle to Low	Intercept	0.307	0.165	3.449	0.063	[a]	
	White to Black/ African American	1.678	0.323	26.992	0.000	5.354	[2.843, 10.081]
	Hispanic to Black/ African American	0.324	0.295	1.202	0.273	1.382	[0.775, 2.466]
	Asian/ Multiracial to Black/ African American	0.625	0.416	2.261	0.133	1.869	[0.827, 4.222]
	White to Hispanic	1.354	0.398	11.558	0.001	3.873	[1.774, 8.454]
	Asian to Hispanic	0.302	0.477	0.400	0.527	1.352	[0.531, 3.441]
	White to Asian/ Multiracial	1.052	0.494	4.531	0.033	2.865	[1.087, 7.551]
	Male to Female	-0.079	0.204	0.151	0.697	0.924	[0.620, 7.551]
Upper Middle to Low	Intercept	-0.085	0.174	0.235	0.627		
	White to Black/ African American	2.975	0.317	88.235	0.000	19.589	[10.530, 36.442]
	Hispanic to Black/ African American	1.374	0.278	24.396	0.000	3.949	[2.290, 6.812]
	Asian/ Multiracial to Black/ African	1.914	0.385	24.769	0.000	6.778	[3.190, 14.401]

Spring Theta Categories		*B*	*SE*	χ^2	*p*	*Exp(B)*	95% CI for *Exp(B)*
	American						
	White to Hispanic	1.601	0.373	18.430	0.000	4.960	[2.388, 10.304]
	Asian to Hispanic	0.540	0.432	1.563	0.211	1.716	[0.736, 4.002]
	White to Asian/ Multiracial	1.061	0.458	5.369	0.020	2.890	[1.178, 7.092]
	Male to Female	0.033	0.201	0.028	0.868	1.034	[0.697, 1.535]
High to Low	Intercept	-2.777	0.363	58.455	0.000		
	White to Black/ African American	4.116	0.437	88.866	0.000	61.296	[26.050, 144.231]
	Hispanic to Black/ African American	1.058	0.561	3.556	0.059	2.880	[0.959, 8.644]
	Asian/ Multiracial to Black/ African American	3.571	0.494	52.151	0.000	35.547	[13.487, 93.691]
	White to Hispanic	3.058	0.551	30.853	0.000	21.287	[7.236, 62.624]
	Asian to Hispanic	2.513	0.597	17.697	0.000	12.345	[3.828, 39.812]
	White to Asian/ Multiracial	0.545	0.483	1.272	0.259	1.724	[0.669, 4.445]
	Male to Female	0.797	0.271	8.636	0.003	2.218	[1.304, 3.773]
Upper Middle to Lower Middle	Intercept	-0.392	0.153	6.572	0.010		
	White to Black/ African American	1.297	0.179	52.228	0.000	3.659	[2.574, 5.202]

Spring Theta Categories		*B*	*SE*	χ^2	*p*	*Exp(B)*	95% CI for *Exp(B)*
	Hispanic to Black/ African American	1.050	0.234	20.184	0.000	2.857	[1.807, 4.517]
	Asian/ Multiracial to Black/ African American	1.288	0.283	20.735	0.000	3.627	[2.083, 6.315]
	White to Hispanic	0.247	0.23	1.162	0.281	1.281	[0.817, 2.008]
	Asian to Hispanic	0.239	0.317	0.566	0.452	1.269	[0.682, 2.363]
	White to Asian/ Multiracial	0.009	0.280	0.001	0.975	1.009	[0.583, 1.745]
	Male to Female	0.113	0.154	0.536	0.464	1.119	[0.828, 1.513]
High to Lower Middle	Intercept	-3.085	0.352	76.971	0.000		
	White to Black/ African American	2.438	0.350	48.452	0.000	11.450	[5.763, 22.747]
	Hispanic to Black/ African American	0.734	0.540	1.844	0.174	2.083	[0.722, 6.007]
	Asian/ Multiracial to Black/ African American	2.946	0.421	49.010	0.000	19.022	[8.339, 43.391]
	White to Hispanic	1.704	0.466	13.383	0.000	5.496	[2.206, 13.696]
	Asian to Hispanic	2.212	0.521	18.036	0.000	9.131	[3.29, 25.342]
	White to Asian/ Multiracial	-0.508	0.320	2.523	0.112	0.602	[0.322, 1.126]
	Male to Female	0.876	0.234	14.005	0.000	2.401	[1.518, 3.798]
High to	Intercept	-2.693	0.351	58.887	0.000		

Spring Theta Categories		*B*	*SE*	χ^2	*p*	*Exp(B)*	95% CI for *Exp(B)*
Upper Middle	White to Black/ African American	1.141	0.344	10.994	0.001	3.129	[1.594, 6.141]
	Hispanic to Black/ African American	-0.316	0.531	0.355	0.552	0.729	[0.258, 2.062]
	Asian/ Multiracial to Black/ African American	1.657	0.389	18.150	0.000	5.245	[2.447, 11.242]
	White to Hispanic	1.457	0.444	10.763	0.001	4.292	[1.798, 10.246]
	Asian to Hispanic	1.973	0.480	16.930	0.000	7.193	[2.810, 18.413]
	White to Asian/ Multiracial	-0.516	0.259	3.983	0.046	0.597	[0.359, 0.991]
	Male to Female	0.763	0.215	12.627	0.000	2.145	[1.408, 3.268]

Note. Multinominal regression was used to predict the difference in the nominal dependent variable, Spring raw scores, given the independent variables of ethnicity and gender.

Results of the multinomial logistic regression analysis for f2f students indicated the odds ratios of White students scoring in a higher category than Black/African American students was statistically significant in all six dependent variable combinations. The greatest of these odds ratios suggested that White students were 61.296, 95% CI [26.050, 144.231], times that of Black/African American students in scoring High versus Low on the spring FastBridge assessment, Wald $\chi^2(1) = 88.866$, $p < .001$. The smallest odds ratio in comparing White to Black/African American students indicated that the former was 3.129, [1.594, 6.141] times more likely to score in the High category opposed to the Upper Middle category, Wald $\chi^2(1) = 10.994$, $p < .01$, than the Black/African American comparison group. White students also had higher odds ratios for scoring in a higher category in five of the six categories when compared to Hispanic students and only one of the six when compared to Asian/Multiracial students. The greatest odds ratio for White to Hispanic students resulted in White students having 21.287 higher odds, 95% CI [7.236, 62.624], Wald $\chi^2(1) = 30.853$, $p < .001$, of scoring in the High group than the Low group. The only category in which the odds ratio for a White student scoring higher than a Hispanic student was not statistically significant was in comparing the Upper Middle group to the Lower Middle group. This may be because most of the students who took the spring assessment scored in these two categories. Finally, the odds of a White student scoring in the Upper Middle category versus the Low category was 2.89, [1.178, 7.092], times that of an Asian/ Multiracial student (Wald $\chi^2(1) = 5.369$, $p < .05$. However, the odds of a White student scoring in the High versus Upper Middle category were about half that of an Asian/Multiracial student, also a statistically significant effect, *Exp(B)* = 0.597, [0.359, 0.991], Wald $\chi^2(1) = 3.983$, $p < .05$, indicating that Asian/Multiracial students were more likely to score High versus Upper Middle when compared to White students.

Asian/Multiracial students had the second most occurrences of statistically significant effects over the other ethnicities with five of six categories of higher odds ratios than Black/African American students and three of six categories of higher odds ratios than Hispanic students. Asian/Multiracial students had higher odds ratios than Black/African American students in every category of the paired dependent variable except Lower Middle to Low. Only 25.8% of Asian/Multiracial students scored in the Lower Middle to Low categories while 69.1% of Black/African American students scored in these two categories. The odds ratio of Asian/Multiracial students scoring in the High versus Low category was 35.547, 95% CI [13.487, 93.691] times that of a Black/African American student, Wald $\chi^2(1) = 52.151$, $p < .001$, representing the greatest odds ratio between the two ethnicities. The smallest statically significant odds ratio indicated that Asian/Multiracial students were 3.627, [2.083, 6.315] times that of Black/African American students to score in the Upper Middle category rather than the Lower Middle (Wald $\chi^2(1) = 20.735$, $p < .001$).

In all three statistically significant ($p < .001$) cases of Asian/Multiracial students having higher odds ratios than Hispanic students, the dependent variable categories were in reference to scoring in the High group as opposed to any other. This suggests that Asian/Multiracial students were more likely to score High than Low (*Exp(B)* = 12.345), Lower Middle (*Exp(B)* = 9.131), or Upper Middle (*Exp(B)* = 7.193) when compared to Hispanic students with 95% CI [3.828, 39.812], [3.29, 25.342], and [2.810, 18.413], respectively.

When considering Hispanic students to the other ethnicities, these students only had a higher likelihood of scoring in higher categories than Black/African American students, and only in the groupings of Upper Middle to Lower Middle and Upper Middle to Low. Hispanic students were 2.857, 95% CI [1.807, 4.517] times more likely to score in the Upper Middle category than Lower Middle category, and 3.949, [2.290, 6.812] times more likely to score in the Upper Middle than Low categories versus Black/African American students. Each had statistically significant effects at $p < .001$, with Wald $\chi^2(1) = 20.184$ and Wald $\chi^2(1) = 24.369$, respectively.

Finally, the results of the multinomial regressions indicated that males were more likely to score High in any other category than females doing the same. Males were more likely to score High opposed to Upper Middle than females, *Exp(B)* = 2.145, 95% CI [1.408, 3.268], Wald $\chi^2(1) = 12.627$, $p < .001$; more likely to score High as opposed to Lower Middle than females, *Exp(B)* = 2.401, [1.518, 3.789], Wald $\chi^2(1) = 14.005$, $p < .001$; and more likely to score High as opposed to Low than females, *Exp(B)* = 2.218, [1.304, 3.773], Wald $\chi^2(1) = 8.636$, $p < .001$.

Due to a lack of data in both the district and state virtual data, ordinal logistic regressions could not be conducted to compare the online programs to the f2f courses, and TDT could not be used as part of the discussions. Descriptive statistics for district virtual ROI information indicated that students either had learning loss or no change from the fall to spring quarters of the 2018-2019 school year, but no other interpretations of standardized assessment success can be made.

Instead, multinomial logistic regressions were conducted as exploratory analyses for the f2f group only, where possible, and data descriptions were substituted when not possible. Overall exploratory analyses for standardized assessment success indicated that, in the f2f group, White and Asian/Multiracial students had higher odds of having larger learning gains and higher FastBridge assessment scores than Black/African American students and Hispanic students. The same held for males having higher overall odds than females. More on this is in Chapter 5.

Analysis for RQ3 – F2F versus District Virtual

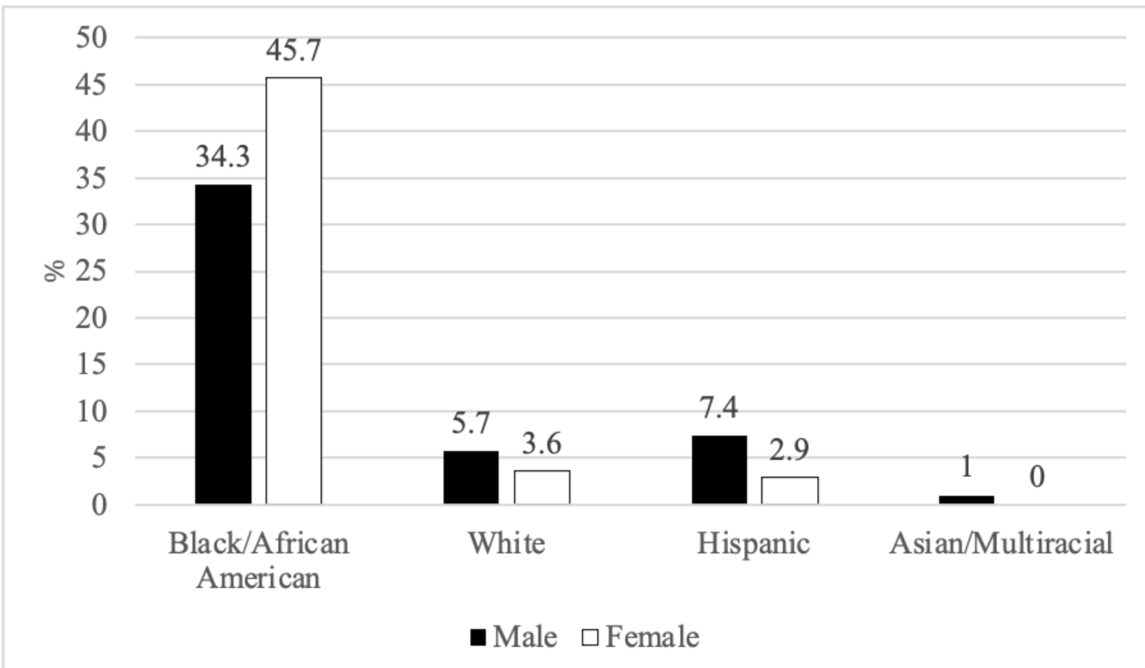
The goal of RQ3 was to identify if there was a difference in whether a student completed or did not complete their Algebra I course, dependent on if they were in a f2f class versus a virtual course. Once data had been filtered to include only information for ninth-grade students, who were first-time Algebra I course takers, 323 students took the courses virtually and 26,547 students who took the course f2f. Of these students, 167 received either an I for incomplete, NG for no grade, or NC for no credit. The data were unclear whether the NG and NC (n = 27) meant that the student had not completed the course or had not received a grade for some other reason, such as a typo or missed grade. Filtering out these two codes left 127 students with incompletes, 0 from the state virtual, 18 from the district virtual courses, and 109 from the f2f courses. After removing NG and NC, of the 26,529 remaining f2f students, 0.41% did not complete, while 5.9% of the remaining 306 district virtual students did not complete. Laerd Statistics states that there must be “a bare minimum of 15 cases per independent variable, although some recommend as high as 50 cases per independent variable” (2017, p. 3). Due to the small number of incompletes for the virtual courses, binary logistic regressions could not be conducted and a final determination as to the differences in course completion for RQ3 could not be made, nor could TDT be used as part of the discussion.

Exploratory Analysis of RQ3 – Gender and Ethnicity Differences for Course Completion

A bar chart representing the percentage of students, both f2f and in the district virtual course, in each category of gender nested within ethnicity was created to try to get an understanding of which students were not completing their courses (see Figure 10).

Figure 10

Percentage of Students Who Did Not Complete the Course per Ethnicity and Gender



Note. Percentages were based on the total number of students who did not complete the Algebra I classes, n = 140.

Summary

The purpose of this quantitative study was to assess and compare learning outcomes for one U.S. region's students who took an Algebra I course in a district-run virtual program, a state-run virtual program, or in a f2f environment to determine the efficacy of the two approaches for moving these students closer to meeting graduation requirements. Archived data was received from the cooperating school district that included information such as gender, ethnicity, course grades, FastBridge growth calculations, FastBridge raw scores, and complete/incomplete codes for students taking Algebra 1 in one of the three designated environments. The data included demographics and course results for students in Grades 9 through 12 who were taking the course for the first time, as a remedial course, and as a credit recovery course, but was filtered to include only ninth-grade students taking Algebra 1 for the first time. Furthermore, while assumption testing was not necessary for census sampling, steps were taken to ensure most assumptions were being met so that results were more reliable.

The results of OLR for RQ1 indicated that course environment, ethnicity, and gender were all contributing factors to the odds of students earning higher grades. End-of-course grades were seemingly higher, on average, for students in f2f courses as opposed to the grades for students in the district virtual courses. Transactional distance may be a contributing factor in the discrepancies between these environments as deficiencies in understanding, motivation, and communication are often greater in online courses (Bergdahl et al., 2020; Usta & Mirasyedioğlu, 2022).

In line with previous research (e.g., Ahn & McEachin, 2017; Hart et al., 2019; Heinrich et al., 2019; Heppen et al., 2017; Hughes et al., 2015), exploratory analysis results indicated that White and Asian/Multiracial students were more likely to earn better grades than Black/African American and Hispanic students and that females had better chances of earning higher grades than males. After looking deeper into the distribution of grades, it was found that Black/African American students accounted for the majority of Fs given by the district and maintained a course grade average of 73.93%. On the other hand, White students had a course grade average of 83.15% and Asian/Multiracial students averaged 80.69%. (Hispanic students averaged 76.33%.) Between gender groupings, females (M = 79.29) had about a 4% higher course average than males (M = 75.84).

Overall course grades had a negatively skewed distribution indicating that more students were passing than failing. However, major spikes in the frequency tables for the f2f and district virtual environments indicated heavy loading right at the 70% mark, the minimum grade needed to pass. These spikes should raise red flags as to how many students were being "bumped up" into passing without necessarily earning it. Several researchers and authors have noted that grade inflation is becoming all too prevalent in high school across the country (Buckley et al., 2018; Chowdhury, 2018; Côté et al., 2020), and the grade distribution noticed with this data seems to echo their findings.

A lack of data for both the state and district virtual courses, in terms of FastBridge assessments, limited the available analysis options for RQ2. Descriptive statistics for district virtual ROI information indicated that students either had learning loss or no change from the fall to spring quarters of the 2018-2019 school year. The two occurrences of learning gains in this group were so minimal that one could speculate they may be attributed to lucky guessing.

In the f2f group, multinomial logistical regressions for the exploratory analyses indicated that White, Asian/Multiracial, and male students had higher odds of having larger learning gains and scoring higher on the FastBridge aMath assessments than Black/African American students, Hispanic, and female students.

Between the results of RQ1 and RQ2, particularly for female and male students, contradictory findings occurred. In RQ1, females tended to outperform male students on end-of-course grades, but the reverse was evident in RQ2. This is discussed further in Chapter 5.

Finally, for RQ3, while ratios for f2f and district virtual courses indicated that a higher proportion of virtual students were not completing their Algebra I course, there was not enough data to conduct any type of true analysis.

Overall lack of data for virtual students, both in the district and state programs, limited the depth to which comparisons between the environments could be made. However, from the data available, prior research, and transactional distance theory, implications for social change can be considered. Based on the understanding of this study’s findings, suggestions for future research are discussed in the coming chapter.

Chapter 5: Discussion

The purpose of my study was to assess and compare learning outcomes for one U.S. region's students who took an Algebra I course in a district-run virtual program, a state-run virtual program, or in a f2f environment to determine the efficacy of the two approaches for moving these students closer to meeting graduation requirements. While research exists that considered the impact that f2f and online learning had on student achievement, very few studies focused on Algebra 1 courses for first-year high school students. Still, students are continuing to be put into online courses with no evidence that the virtual environment is as effective as f2f courses in promoting academic achievement. The goal of conducting this study was to try to fill a gap in the understanding of how well students performed in online courses compared to f2f.

I used ordinal logistic regressions to decipher the difference between the district virtual and f2f courses for student achievement based on end-of-course grades. Binomial and multinomial logistic regressions were used when possible to evaluate the differences in learning gains based on FastBridge ROI and the 2018-2019 FastBridge spring assessment scores. When too few data points were present, as was the case for the state virtual program, descriptions, tables, and figures were provided to represent the data as thoroughly as possible.

Overall findings of this quantitative study suggested that students in the f2f courses had higher course grades and higher odds of getting better grades than students in virtual courses. Transactional distance was used to explain the differences in end-of-course grades for RQ1. A lack of data and inability to analyze f2f and online courses for RQ2 and RQ3 limited possible discussions of the impact of transactional distance on FastBridge aMath assessments and completion rates.

A deeper look into the breakdown of grades and scores between ethnicities and genders through exploratory analyses revealed that Black/African American and Hispanic students were underperforming compared to their White and Asian/Multiracial classmates, as were males to females. In nearly all cases measured, White and Asian/Multiracial students had better odds of scoring in higher categories of end-of-course grades, ROI, and FastBridge raw scores than Black/African American students, and, with smaller disparities, better than Hispanic students. In cases where only descriptive statistics could be completed, Black/African American students consistently had some of the lowest achievement averages while White students carried some of the highest averages. In terms of gender, females outperformed males on course grades and ROI, but males had higher odds of scoring higher on the spring assessment.

Interpretation of the Findings

Previous research regarding the comparisons of online versus f2f education has resulted in mixed reports as to the effect of course environment on academic achievement (see Ahn & McEachin, 2017; Chisadza et al., 2021; Hart et al., 2019; Heinrich et al., 2019; Heppen et al., 2017; Hughes et al., 2015). Many of the older studies on f2f and online education tended to indicate that online learning allows for more flexibility and self-paced learning than f2f, but effect sizes were mostly insignificant or small indicating no difference in student learning between the two environments (Arias et al., 2018;

Chisadza et al., 2021; Hart et al., 2019; Hughes et al., 2015). Others suggested that students, especially younger, early secondary students, were not equipped with the self-discipline, motivation, or know-how to successfully pass an online course as they might in a f2f class (Bergdahl et al., 2020; Heinrich et al., 2019; Kontkanen et al., 2017; Larkin & Jamieson-Proctor, 2015; Vickrey et al., 2018). More recent studies tended to vary in findings dependent on the course topic, age group, location, prior GPA., learning styles, and gender (Amparo et al., 2018; Brubacher & Silinda, 2019; Cheng et al., 2017; Owan & Asuquo, 2021). Consistent with Amparo et al. (2018), Heppen et al., 2017, and others, the results of this study seemed to indicate that f2f students had higher end-of-course grades than online students. However, a lack of data, especially for ROI and FastBridge assessments, made it impossible to say for certain which environment was more conducive to learning gains.

As noted in Chapter 4, the results of this study seemed to emulate those of Ahn and McEachin (2017), Heinrich et al. (2019), Heppen et al. (2017), Heissel (2016), and Wakil et al. (2019) in that they point to the understanding that online students have lower test scores, lower GPAs, and lower overall academic gains. Specifically, students in the district online courses had an 8% lower average end-of-course score, nearly one full letter grade, than the average for the f2f students. Unfortunately, comparisons between the online and f2f environments could only be established for the course grades due to limited data. As noted previously, these results must be interpreted with caution. While longitudinal archived data and census sampling can help alleviate validity issues, the lack of data and possible sampling bias does not allow for definitive conclusions to be made as to whether the course environment was the main attributor to the difference in grades or if other extraneous variables were at play. Comparisons between f2f and online courses could not be made in terms of standardized assessment success or course completion.

Heppen et al. (2017) and Hughes et al. (2015) both suggested that the students in the online courses that were not performing as well as the f2f courses were mostly Black or Hispanic students and students that qualify for free or reduced lunches. The data for this study did not have enough cases to speak to the trends of Black/African American or Hispanic students' online performance, but their findings were reflected in the overall outcomes of this study. Black/African American students made up just about half of the population (48.9%), while White students made up about one-third (28.9%) and Hispanic students less than that (15.6%). Yet, Black/African American students accounted for almost three-quarters (70.7%) of all Fs in the district and less than one-third (27.8%) of the As, while White students accounted for almost half of the As (47.2%) and less than 10% of the Fs (8.8%). Hispanic students' contribution to Fs and As stayed consistent with their population proportions at about 15% each. The findings for course grades within this research were also mostly consistent with those found for ROI and the spring FastBridge raw assessment scores in that White students were performing better and had higher odds of performing better than Black/African American students. In some cases, the odds of White students scoring higher on the assessments were fairly extreme compared to their Black/African American counterparts. Easton et al. (2017) and Paschall et al., (2018) also found that White and Asian students earned higher grades than

Black/African American and Hispanic students. Kuhfeld et al. (2018) found comparable results in their study of ethnic and socioeconomic achievement gaps and wrote “Within students who are not in poverty, White students almost always outperformed Black and Hispanic students…[and] White students in poverty generally performed similarly to Black and Hispanic students who are not in poverty” (p. 68). In other words, White students were generally outperforming Black and Hispanic students in most situations.

Ahn and McEachin (2017), in studying online and f2f learning in North Carolina, found that online students were mostly White, mostly female, were less likely to pass the end-of-course assessment, and were making fewer learning gains. Partially contradictory to these findings, in terms of the gender gaps, this research found that females had higher odds of earning better grades than males and were more likely to have larger learning gains than males, particularly in having large learning gains as opposed to minimal learning gains or loss. However, like Ahn and McEachin, male students had higher odds of scoring higher on the spring assessment. The discrepancies between who was performing better on the ROI and assessment categories would make sense if males were also scoring higher on the fall FastBridge assessments. Consistent higher scores on the fall assessment and spring assessments meant the difference between the two would be smaller, thus smaller learning gains. This also suggested that assessment scores between males and females in 10th grade should be less differential if females were making bigger gains, but that was beyond the scope of this study.

Sutton et al. (2018) found that during the transition from eighth to ninth grade, female students tended to maintain their GPAs from eighth grade while male students’

GPAs, on average, tended to have a statistically significant drop during ninth grade. Sutton et al. (2018) found that between ethnicity and gender groupings, the greatest differences were found between White females and Black/African American males, while Black/African American females to males, in favor of females, had the greatest differences within the same ethnicity. The results of the present research, the observed average drops in GPA by Sutton et al. (2018) and findings by Heppen et al. (2017), Kuhfeld et al. (2018), Easton et al., (2017), Paschall et al., (2018), and Hughes et al. (2015) all provided an explanation as to what was happening as far as the differences between ethnicities and gender achievement gaps, especially for course grades and GPA, but did not provide insight as to the differences in assessment scores or why these gaps exist.

A look into the breakdown of the course grades showed nonnormal distribution curves for White students, Black/African American students, and overall district. In my study, in all environments, there was a significant difference in the frequency in which students earned a 69%, which was failing, and a 70%, which was passing, presenting red flags as to possible grade inflation. Black/African American students accounted for 65.4% of the "just passing" mark. Chowdhury (2018) noted that grade inflation was the giving of higher grades without evidence of higher achievement, and that grade compression occurs when a normal distribution curve was skewed so that more students were at the top of the curve, often because of grade inflation. Chowdhury also suggested that grade inflation may occur when the educational system was trying to "maintain a positive public image" (p. 88), teachers were concerned about their students' psychology (especially when parental and social pressures were involved), teachers' pay was affected by having higher pass rates, and when teachers were trying to save time "because they do not want to spend most of their office hours giving justifications for low grades on assignments or examinations" (p. 87). Sorubakhsh-Castillo (2018) and Côté et al. (2020) stated that grade inflation may occur when coaches and administrators were pressuring teachers to pass students who would not be able to pass on their own. While it was unclear whether grade inflation was truly happening in this school district, the frequency in which students earned a grade of 70% versus a grade of 69% was cause for concern, especially amongst the Black/African American students. "There is an innate obligation instilled in academic institutions to inflate grades to help their poorer-performing students to acquire better opportunities" (Chowdhury, 2018, p. 91), however it is unclear whether this is beneficial to the student. A vast number of employers, university faculty, journalists, and statisticians believed first-year college and university students were not ready for college-level work (e.g., Butrymowicz, 2017; Chen, 2022; Côté et al., 2020; Grace-Odeleye & Santiago, 2019; Marcus, 2016; Ostashevsky, 2016; Villares & Brigman, 2019).

Regarding transactional distance, although the TDT framework has been thoroughly established as being valid when considering online learning, a lack of data for the online courses prohibited the use of applying TDT to understand the differences between online and f2f courses for standardized assessment success and course completion. According to prior research, students in a f2f course should have had minimal transactional distance when compared to students in the virtual courses. This may, in part, explain why students in the f2f courses seemed to outperform students in the online courses. However, with over 98% of students being in the f2f courses, it was not possible to conclude with certainty as to whether transactional distance had a role in students' learning achievements or if there were other factors at play.

Limitations

Research studies were often limited by cost, time, available participants and/or data, and the strength of underlying assumptions (Gopalan et al., 2020). No study can truly account for every possible variant in participant differences, study environments, or other influences which may affect the behavior of an independent variable on the dependent variable (Gelman & Imbens, 2019; Gopalan et al., 2020). As such, researchers should identify the realms under which their studies have limitations.

This study was primarily limited by its design. Gopalan et al. (2020) stated that quasi-experimental designs, while able to determine causal effects, are not able to decipher how or why changes in the dependent variable occur. One of the greatest limitations of this design was the lack of entry-level data to determine group comparability prior to conducting analyses. Nonequivalent quasi-experimental posttest only research has been conducted in the social sciences in recent years (Zb et al., 2021), but all were limited in similar capacities as this study inability to precisely ascertain if the independent variable was the cause for changes in the dependent variable (Dawson, 1997). While elements outside of the control of this study could have affected results, and most likely did, (Cole et al., 2021; Martin & Bolliger, 2018; Valverde-Berrocoso et al., 2020) this study, like many others in educational research, has validity issues and should be interpreted with caution (see Fryer & Bovee, 2018; Niu, 2020; Valverde-Berrocoso et al., 2020).

Additionally, the findings of these types of studies are not generalizable to populations outside of the specific population used in the study (Gopalan et al., 2020). The population of this study was that of ninth-grade students taking Algebra I for the first time, either online or f2f, in one school district in the United States. While the present findings may indicate that f2f students, White students, and female students tended to perform better than online, Black/African American, or male students, this does not mean the same will hold true for students in other states, countries, or even other grades within the same district. Readers of this study should understand that the findings and implications made here must be considered with caution, especially when standing alone. However, great care was taken to include past literature, both in agreement and opposition, that aids in strengthening the results discussed.

Another limitation of the study lies in the use of the selected independent variables, and how students were placed in courses. Whether it be due to student preference, scheduling conflicts, health reasons, administrative placements, or other reasons, it was assumed that students were not placed strictly by chance. The inconsistent method by which students were placed into the three types of classrooms was considered a limitation because it restricts the possibility of random assignment of students to courses and the ability to balance course group size to make comparisons between the groups more reliable.

The use of student course grades as a dependent variable was potentially hazardous as these grades were generally subjective scores that were not standardized and most likely included some measurement errors (Chowdhury, 2018; Côté et al., 2020; Sorubakhsh-Castillo, 2018). To circumvent results being solely based on course grades, data was requested for the state standardized assessment scores which would have proved more reliable. However, with the adoption of a newer assessment program, scores were sent for the most recent year only, meaning many students did not yet have scores recorded. Of the 26,887 Algebra I students in ninth grade, 1,061 students had ROI and/or spring assessment scores, and only nine of those were from the virtual courses. The limited data simply did not allow for in-depth comparative analysis.

Several educators and researchers claimed there were several f2f and virtual programs that either lowered their standards for scope and rigor so students could receive a passing grade or were themselves limited in their ability to maintain academic honesty amongst students (Lucky, et al., 2019; Powell et al., 2015). This study was limited in that it did not present any judgments one way or another as to whether the online courses involved were equitable to the f2f courses in maintaining standards and providing genuine learning opportunities through meaningful interactions with peers, teachers, and content. Additionally, this study did not attempt to make comparisons regarding the depth and breadth of the covered content between the three environments, or to what extent teachers may have influenced student achievement. Teacher interactions could be in the form of curving grades to better create a normal distribution, pedagogy, such as allowing test or quiz retakes in some courses but not in others, or grade inflation (Calsamiglia &

Linglio, 2019; Chowdhury, 2018; Côté et al., 2020; Lu et al., 2021). Without a deeper understanding of how the three environments compared as far as content and rigor, this study was strictly limited to supplying an informational foundation upon which further studies can be conducted.

Finally, several studies have mentioned that student demographics, such as ethnicity and socioeconomic status, can play a role in predicting student achievement (e.g., Cavanaugh & Jacquemin, 2015; Cooper et al., 2019; Easton et al., 2017; Hart et al., 2019; Heinrich et al., 2019; Kuhfeld et al., 2018; Paschall et al., 2018; Protopsaltis & Baum, 2019). Unfortunately for this study, socioeconomic status was not an available data point, and therefore could not be used as a covariate. This limited the study in that it was impossible to specify whether student performance was affected by low or high family income.

The lack of data available for this study and its many limitations should leave the reader using its findings with caution. The findings of this study should not be used as a clear indication as to whether online or f2f learning had greater effects on achievement for ninth-grade students taking Algebra 1 for the first time. However, the study can be used as a base on which to build further studies comparing the effectiveness of online and f2f courses.

Recommendations for Future Studies

Consistent with past research, the findings of this study suggested that White students were outperforming Black/African American students in grades, learning gains, and test scores, and that females were outperforming males in all categories except test scores. However, it was unclear why these discrepancies were occurring. Future research should look into the mechanisms at play that are causing continued differences between ethnicities and genders, such as test anxiety, perceptions of the importance of school, and the extracurricular activities in promoting "grade bumps".

Due to the abundant amount of f2f data and the insufficient amount of data for virtual learning, the question of the availability of quality records in sufficient amounts comes to mind. If school districts are spending the money to ensure virtual courses are available for students, either as general education, remedial courses, or advanced placement, then it is presumed they would also maintain detailed records for the students of these courses to confirm they are spending their money where it should be spent. It would seem that record keeping for online courses should be easier than for f2f courses since all of the data is already online. However, this did not seem to be the case in many instances. One suggestion for future research would be to look into the discrepancies between f2f and virtual recordkeeping, and how to improve the available data for virtual students.

The lack of standardized assessment data was most likely due to the introduction of the newer FastBridge assessments as opposed to the Milestones. As such, the narrow comparisons between the virtual courses and f2f courses were made based on one school year's assessment scores, and only from a small fraction of students. Once more students have taken the FastBridge assessments and more data becomes available, a follow-up study should be conducted to determine if the trends in odds ratios remain consistent with the findings of this study. The lack of data also affected the degree to which TDT could be applied to the findings. Future research, utilizing the data for the year provided, could be conducted as an exploratory analysis using TDT as a basis, perhaps as a single-case analysis.

The discrepancies in FastBridge ROI and spring raw scores for males and females may suggest a difference in how males and females approach educational goals. Being that RQ1 was representative of a full year's worth of work and RQ2 was representative of a few hours' worth of testing time, it may be the case that females were more dedicated to long-term achievement but did not perform well on standardized tests. On the other hand, male students may have performed better on short achievement tasks as opposed to those drawn out over a semester or year. Future research on educational goals and goal attainment between genders, particularly for high school students, could provide exceptionally valuable information on pedagogies when teaching the two groups.

Evaluating the comparability of content and rigor of online Algebra I courses was beyond the scope of this study. Echoing the recommendations of several other researchers, future research should attempt to compare online and f2f courses based on pedagogy, teaching materials used, content covered, and grading policies. Additionally, the data used in this study were from the school years 2016-2017 to 2018-2019, pre-Covid-19. It would be interesting to see if the trends found here and in past research continue in a postCovid-19 era.

Finally, the original proposal for this study was to compare online and f2f outcomes for high school geometry credit recovery students in a school district on the east coast of the United States. However, a fire in one of the high schools and the onset of

Covid-19 caused the original cooperating school district to back out of the research. Once a new cooperating school district was found, the districts' interests changed the focus of the study to be on ninth-grade Algebra I first-time course-takers. Future research should consider the differences in learning outcomes for Algebra I credit recovery students, geometry first-time course-takers, and geometry credit recovery students.

Implications for Social Change

In studying the completion and success of high school Algebra I students in online and f2f environments, the intention of this study was to learn more about the achievement of students in online and f2f courses and to possibly provide more insight for school administrators and counselors to place students in an environment best suited to individual needs. However, due to a lack of available data for virtual students, comparisons between the environments could not be conducted except in the case of course grades. As noted in the recommendations section, future research should address the issue of record keeping for online courses. An unintended implication for social change resulting from this study was that it highlighted how incomplete archived data sets could be and how prohibitive they could be for in-depth data analysis. If administrators and, to an extent, policymakers genuinely want to make data-driven informed decisions as to the future of education, they need to start by looking into gaps in their own data sets.

Exploratory analysis results add to the growing number of studies with results indicating a difference in performance between Black/African American, Hispanic, and White students, and males and females. Additionally, even though the findings of this study could not be linked with TDT, the in-depth literature review may encourage others to take a closer look at how TDT could be used as a framework for future studies comparing online and f2f environments for high school math students.

Conclusion

Students everywhere are facing an ever-growing demand to graduate from high school, to gain acceptance into postsecondary institutions, or to increase their chances of higher-paying career opportunities (Bureau of Labor Statistics, 2017a, 2017b). For those that don’t graduate, dropping out of high school affects not only the student’s future opportunities for education and employment but also their means of living and quality of life. As a means of meeting the demand for getting students to pass required courses and graduate from high school, many school divisions across the nation have begun offering high school credits online, including Algebra I courses for first-year high school students. Online courses are being offered to allow students the flexibility to earn credits by being able to work from anywhere and at any time, potentially making the courses shorter and more individualized than a f2f course. However, there was little to no evidence that online courses were as effective as their f2f counterparts in getting students closer to earning credits for graduation for this school district.

Ordinal and multinomial logistic regression calculations were used to analyze archived data collected through census sampling to assess and compare learning outcomes for one U.S. region’s students who took an Algebra I course in a district-run virtual program, a state-run virtual program, or in a f2f environment. A lack of data for virtual students forced calculations to be curtailed to mostly descriptions of the available cases, except for analyzing course success based on course grades. F2f data was explored in depth for both course grades and spring assessment scores, resulting in observed differences in ethnicity and gender achievements. Future research must look more closely into the mechanisms contributing to the achievement gaps if evidence-based educational reform is ever going to make strides in closing these inequalities.

References

Abuhassna, H., Al-Rahmi, W. M., Yahya, N., Zakaria, M. A. Z. M., Kosnin, A. B. M., & Darwish, M. (2020). Development of a new model on utilizing learning platforms to improve students' academic achievements and satisfaction. *International Journal of Educational Technology in Higher Education*, *17*(1). https://doi.org/10.1186/s41239-020-00216-z

Achieve. (2016, November 7). *State-by-state graduation requirements (Class of 2015)*. https://www.achieve.org/publications/state-state-graduation-requirements-class-2015

Adams, S. (2020). *A study of at-risk high school students enrolled in a full-time online learning environment*. [Doctoral dissertation, City University of Seattle]. ProQuest Dissertations Publishing. https://www.proquest.com/dissertations-theses/study-at-risk-high-school-students-enrolled-full/docview/2505407482/se-2

Adhikari, G. P. (2021). Calculating the sample size of quantitative studies. *Scholars' Journal*, *4*(1), 14-29. https://www.nepjol.info/index.php/scholars/article/view/42458/32261

Ahmad, H. B., & Halim, H. B. (2017). Determining sample size for research activities: the case of organizational research. *Selangor Business Review*, *2*(1), 20-34. https://sbr.journals.unisel.edu.my/ojs/index.php/sbr/article/view/12/20

Ahn, J., & McEachin, A. (2017). Student enrollment patterns and achievements in Ohio's online charter schools. *Educational Researcher*, *46*(1), 44-57. https://doi.org/10.3102/0013189X17692999

Akos, P., & James, S. (2020). Are course withdrawals a useful student success strategy? *National Academic Advising Association Journal, 40*(1), 80-93. https://doi.org/10.12930/NACADA-18-34

Albano, A. D., Christ, T. J., & Cai, L. (2018). Evaluating equating in progress monitoring measures using multilevel modeling. *Measurement: Interdisciplinary Research and Perspectives, 16(*3), 168-180.

Alfrisa, A. M., A'yunna, D. Q., & Permatasari, W. A. (2020). The effectiveness of using blended learning in mathematics teaching on algebra. *PRISMA, Prosiding Seminar Nasional Matematika, 3,* 317-321.

Ali, A., Ali, S., Khan, S. A., Khan, D. M., Abba, K., Khalil, A., Manzoor, S., & Khalil, U. (2019). Sample size in multilevel logistic regression models. *PloS One, 14(*11), e0225427. http://doi.org/10.1371/journal.pone.0225427

Aluko, F. R., & Omidire, M. F. (2021). A critical review of student assessment practices in distance education in an emerging economy: Benchmarking practices against policy. *African Education Review, 17*(5), 76-94.

Alvarez & Marsal. (2018). *Final report District of Columbia public schools audit and investigation* (Contract number CW57247). Office of the State Superintendent of Education. https://osse.dc.gov/sites/default/files/dc/sites/osse/release_content/attachments/Report%20on%20DCPS%20Graduation%20and%20Attendance%20Outcomes%20-%20Alvarez%26Marsal.pdf

American Institute for Research. (2017, June*). Getting back on track: Course progression for students who fail Algebra I in ninth grade (Research Brief 6)*.

Amparo, A. R., Smith, G., & Friedman, A. (2018). Gender and persistent grade performance differences between online and f2f undergraduate classes. In T. Bastiaens, J. Van Braak, M. Brown, L. Cantoni, M. Castro, R. Christensen, G. Davidson-Shivers, K. DePryck, M. Ebner, M. Fominykh, C. Fulford, S. Hatzipanagos, G. Knezek, K. Kreijns, G. Marks, E. Sointu, E. Korsgaard Sorensen, J. Viteli, J. Voogt, P. Weber, E. Weippl & O. Zawacki-Richter (Eds.), *Proceedings of EdMedia: World Conference on Educational Media and Technology* (pp. 1935-1939). Association for the Advancement of Computing in Education.

Anderson, N. P. (2016). Plato online credit recovery programs for meeting promotional and graduation requirements. *Distance Learning, 13*(2), 7-12.

Arias, J. J., Swinton, J., & Anderson, K. (2018). Online vs. f2f: A comparison of student outcomes with random assignment. *E-Journal of Business Education and Scholarship of Teaching, 12*(2), 1-23.

Atun, H., & Usta, E. (2019). The effects of programming education planned with TPACK framework on learning outcomes. *Participatory Education Research, 6*(2), 26-36. http://dx.doi.org/10.17275/per.19.10.6.2

Atwell, M., Balfanz, R., Manspile, E., Byrnes, V., & Bridgeland, J. (2021). *Building a grad nation: Progress and challenge in raising high school graduation rates*. https://files.eric.ed.gov/fulltext/ED617355.pdf

Austin, P. C., & Steyerberg, E. W. (2017). Events per variable (EPV) and the relative performance of different strategies for estimating the out-of-sample validity of logistic regression models. *Statistical Methods in Medical Research, 26*(2), 796-808. https://doi.org/10.1177/0962280214558972

Barbour, M. K. (2017). *Virtual Schools in the U.S 2017*. National Education Policy Center. https://nepc.colorado.edu/sites/default/files/publications/Section%20II%20Virtual%20Schools%202017.pdf

Barnett, K. K. (2016). The at-risk student's journey with online course credit: Looking at perceptions of care. *Journal of Online Learning Research*, *2*(4), 367-398.

Bartlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organizational research: Determining appropriate sample size in survey research. *Information Technology, Learning, and Performance Journal*, *19*(1), 43-50. https://www.opalco.com/wp-content/uploads/2014/10/Reading-Sample-Size1.pdf

Behi, R., & Nolan, M. (2014). Causality and control: Threats to internal validity. *British Journal of Nursing, 5*(6).

Belland, B. R., Weiss, D. M., Kim, N. J., Piland, J., & Gu, J. (2019). An examination of credit recovery students' use of computer-based scaffolding in a problem-based, scientific inquiry unit. *International Journal of Science and Mathematics Education*, *17*(1), 273-293.

Benson, R., & Samarawickrema, G. (2009). Addressing the context of e-learning: Using transactional distance theory to inform design. *Distance Education, 30*(1), 5-21.

Bergdahl, N., Nourni, J., Fors, U., & Knutsson, O. (2020). Engagement, disengagement and performance when learning with technologies in upper secondary school. *Computers and Education, 149*, 103783.

Bielinski, J., Brown, R., & Wagner, K. (2020, August). *COVID Slide: Research on learning loss and recommendations to close the gap.* Illuminate Education. https://f.hubspotusercontent20.net/hubfs/5196620/covid-19-slide-whitepaper.pdf

Bielinski, J., Brown, R., & Wagner, K. (2021, February). *No longer a prediction: What new data tell us about the effects of 2020 learning disruptions.* Illuminate Education. https://www.illuminateed.com/wp-content/uploads/2021/06/LearningLoss_2021_Update_FB_031221.pdf

Biery, J. L. (2022). *The relationship between an online mindfulness intervention and academic achievement, attendance, behavior, and social-emotional growth in ninth grade students.* [Doctoral dissertation, Kansas State University]. K-Rex.

Bollinger, D. U., & Halupa, C. (2018). Online student perceptions of engagement, transactional distance, and outcomes. *Distance Education, 39*(3), 299-316.

Boyle, F., Kwon, J., Ross, C., & Simpson, O. (2010). Student-student mentoring for retention and engagement in distance education. *Open Learning, 25*(2), 115-130.

Brown, R. (2021). *FastBridge Learning: Understanding student growth data*. https://www.fastbridge.org/understanding-student-growth-data-0418ascd/download-understanding-student-growth-data-0418ascd/

Brubacher, M. R., & Silinda, F. T. (2019). Enjoyment and not competence predicts academic persistence for education students. *International Review of Research in*

Open and Distributed Learning, *20*(3), 165-179. https://www.irrodl.org/index.php/irrodl/article/view/4325/5140

Buckley, J., Letukas, L., & Wildavsky, B. (Eds.). (2018). *Measuring success: Testing, grades, and the future of college admissions*. Johns Hopkins University Press.

Bureau of Labor Statistics. (2017a). *College enrollment and work activity of 2016 high school graduates (United States Department of Labor No. 17-0477)*. https://www.bls.gov/news.release/pdf/hsgec.pdf

Bureau of Labor Statistics. (2017b). *Unemployment rates and earning by educational attainment*. https://www.bls.gov/emp/ep_table_001.htm

Bureau of Labor Statistics. (2020). *Occupational outlook handbook: Occupation finder*. https://www.bls.gov/ooh/occupation-finder.htm?pay=&education=&training=Short-term+on-the-job+training&newjobs=&growth=&submit=GO

Bush, S. (2015). Sample size determination for logistic regression: A simulation study. *Communications in Statistics – Simulation and Computation*, *44*(2), 360-373.

Butrymowicz, S. (2017, January 30). *Most colleges enroll many students who aren't prepared for higher education*. Hechinger Report. https://hechingerreport.org/colleges-enroll-students-arent-prepared-higher-education/

Buzzai, C., Sorrenti, L., Costa, S., Toffle, M. E., & Filippello, P. (2021). The relationship between school-basic psychological need satisfaction and frustration, academic engagement and academic achievement. *School Psychology International, 42*(5), 497-519.

Calsamiglia, C., & Linglio, A. (2019). Grading on a curve: When having good peer is not good. Economics of Education Review, 73, 1-21. https://reader.elsevier.com/reader/sd/pii/S0272775718306174?token=A83EFB8F3D46A1720C86BD9A5DF141ABC0A793E5CB632290FF67170E8FABD2C12080E1D2F1A7670ED450D73FFF14CEE4&originRegion=us-east-1&originCreation=20221014192305

Carr, E. L. (2021). *An investigation into African-American males' achievement growth in reading and math, attendance, and discipline based on after-school program attendance* (Order No. 28323183). [Doctoral Dissertation, Trevecca Nazarene University]. Available from ProQuest One Academic. (2551215166). https://www.proquest.com/dissertations-theses/investigation-into-african-american-males/docview/2551215166/se-2

Cavanaugh, C., Hargis, J., & Mayberry, J. (2016). Participation in the virtual environment of blended college courses: An activity study of student performance. *International Review of Research in Open and Distributed Learning, 17*(3). https://www.irrodl.org/index.php/irrodl/article/view/1811/3760

Cavanaugh, J. K., & Jacquemin, S. J. (2015). A large scale comparison of grade based student learning outcomes in online vs. f2f courses. *Online Learning, 19(*2). https://olj.onlinelearningconsortium.org/index.php/olj/article/view/454/138

Cedar Tutoring Academy. (2018). Why is algebra important? The reasons why it is vital to learn. https://cedartutoring.com/uncategorized/why-is-algebra-important-the-reasons-why-it-is-vital-to-learn/

Center on Innovations in Learning. (2015, January 22). Online credit recovery school program. http://www.centeril.org/reports/resources/2015_01.22OnlineCreditRecoverySchoolProgram.pdf

Chapman, C., Laird, J., Ifill, N., & Kewal Ramani, A. (2011). *Trends in High School Dropout and Completion Rates in the United States: 1972 – 2009* (NCES 2012-006). U.S. Department of Education. Washington, D.C.: National Center for Education Statistics. http://nces.ed.gov/pubsearch

Chen, G. (2022, February 12). Are high school graduate ready for college? Studies are dismal. Public School Review. https://www.publicschoolreview.com/blog/are-high-school-graduates-ready-for-college-studies-are-dismal

Chen, Y., & Willits, F. K. (1999). Dimensions of educational transactions in a videoconferencing learning environment. *American Journal of Distance Education, 13*(1), 45-59. https://eric.ed.gov/?id=EJ588218

Cheng, F. F., Chiu, C. C., Wu, C. S., & Tsaih, D. C. (2017). The influence of learning style on satisfaction and learning effectiveness in the asynchronous web-based learning system. *Library Hi Tech, 35*(4), 473-489. https://doi.org/10.1108/LHT-12-2016-0151

Child Care and Early Education Research Connections. (2019). *Experiments and quasi-experiments* (Project No. 90YE250) [Grant]. U.S. Department of Health and Human Services (HHS), Administration for Children and Families, Office of Planning, Research and Evaluation, Research Connections. https://www.researchconnections.org/childcare/datamethods/experimentsquasi.jsp

Chisadza, C., Clance, M., Mthembu, T., Nicholls, N., & Yitbarek, E. (2021). Online and f2f learning: Evidence from students' performance during the Covid-19 pandemic. African Development Review, 33(1), pp. 114-125. https://doi.org/10.1111/1467-8268.12520

Choi, J., Walters, A., & Hoge, P. (2017). Self-reflection and math performance in an online learning environment. *Online Learning Journal*, *21*(4), 79-102. https://doi.org/10.24059/olj.v21i4.1249

Chowdhury, F. (2018). Grade inflation: causes, consequences and cure. Journal of Education and Learning, 7(6), 86-92. https://doi.org/10.5539/jel.v7n6p86

Clements, M., Stafford, E., Pazzaglia, A. M., & Jacobs, P. (2015). *Online course use in Iowa and Wisconsin public high schools: The results of two statewide surveys* (REL 2015-065). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. https://ies.ed.gov/ncee/rel/regions/midwest/pdf/REL_2015065.pdf

Clements, M., Zweig, J., & Pazzaglia, A. M. (2015). Stated Briefly: Online course use in New York high schools: Results from a survey in the Greater Capital Region

(REL 2015-074). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. http://ies.ed.gov

Coalition for Evidence-Based Policy. (2010). Checklist for reviewing a randomized controlled trial of a social program or project, to assess whether it produced valid evidence. http://coalition4evidence.org/wp-content/uploads/2010/02/Checklist-For-Reviewing-a-RCT-Jan10.pdf

Coalition for Evidence-Based Policy. (2014a). Which study designs are capable of producing valid evidence about a program's effectiveness? A brief overview. http://coalition4evidence.org/wp-content/uploads/2014/10/Which-Study-Designs-are-Capable-of-Producing-Valid-Evidence-of-Effectiveness.pdf

Coalition for Evidence-Based Policy. (2014b). Which comparison-group ("quasi-experimental") study designs are most likely to produce valid estimates of a program's impact?: A brief overview and sample review form. http://coalition4evidence.org/wp-content/uploads/2014/01/Validity-of-comparison-group-designs-updated-January-2014.pdf

Cole, A. W., Lennon, L., & Weber, N. L. (2021). Student perceptions of online active learning practices and online learning climate predict online course engagement. *Interactive Learning Environments, 29*(5), 866-880. https://doi.org/10.1080/10494820.2019.1619593

Conway, K., Hachey, A., & Wladis, C. (2016). Assessing readiness for online education: Research models for identifying students at risk. *Online Learning, 20*(3), 97-109. https://files.eric.ed.gov/fulltext/EJ1113351.pdf

Cooper, E. E., McGee, J. R., Levine-Brown, P., & Bolt, L. (2019). The effectiveness of redesigns in developmental education. *Journal of Developmental Education, 43*(1), 20-32. https://www.proquest.com/scholarly-journals/effectiveness-redesigns-developmental-education/docview/2317570374/se-2

Côté, J., Grayson, J. P., Roberts, S., & Chen, L. H. (2020). Beyond HEQCO's skills assessments. Canadian Journal of Educational Administration and Policy, 194. https://journalhosting.ucalgary.ca/index.php/cjeap/article/view/69016

Cuncic, A. (2019, September 02). Understanding internal and external validity. https://www.verywellmind.com/internal-and-external-validity-4584479

Curtis, H., & Werth, L. (2015). Fostering student success and engagement in a K-12 online school. *Journal of Online Learning Research, 1*(2), 163-190. https://files.eric.ed.gov/fulltext/EJ1148836.pdf

Darling-Aduana, J. (2019). Behavioral engagement shifts among at-risk high school students enrolled in online courses. *AERA Open, 5*(4). https://doi.org/10.1177/2332858419887736

Dawson, T. E. (1997, January 23-25). *A primer on experimental and quasi-experimental design.* [Paper presentation]. Annual Meeting of the Southwest Educational Research Association, Austin, TX, United States.

DeTure, M. (2004). Cognitive style and self-efficacy: Predicting student success in online distance education. *American Journal of Distance Education, 18*(1), 21-38. https://doi.org/10.1207/s15389286ajde1801_3

Dewey, J. (1939). *John Dewey's philosophy*. J. Ratner (ed.). New York: The Modern Library.

Dhivyadeepa, E. (2015). *Sampling techniques in educational research.* Laxmi Book Publication.

Dikmen, C. H., & Demirer, V. (2022). The role of technological pedagogical content knowledge and social cognitive variables in teachers' technology integration behaviors. *Participatory Educational Research, 9(*2), 398-415. http://doi.org/10.17275/per.22.46.9.2

District Virtual. (2021). About. https://www.fultonschools.org/Page/694

Dockter, J. (2016). The problem of teaching presence in transactional theories of distance education. *Computer and Composition, 40*, 73-86. https://www.sciencedirect.com/science/article/abs/pii/S8755461516300305?via%3Dihub

Dorn, E., Hancock, B., Sarakatsannis, J., & Viruleg, E. (2020). COVID-19 and student learning in the United States: The hurt could last a lifetime. *McKinsey & Company, 1*. https://www.childrensinstitute.net/sites/default/files/documents/COVID-19-and-student-learning-in-the-United-States_FINAL.pdf

Duc, T. H. (2012). *Designing distance learning for the 21st century: Constructivism, Moore's transactional theory and web 2.0* (Unpublished master's thesis). Blekinge Institute of Technology, Sweden. http://urn.kb.se/resolve?urn=urn:nbn:se:bth-5707

Duckworth, A., & Gross, J. J. (2014). Self-control and grit: Related but separable determinants of success. *Current Directions in Psychological Science, 23*(5), 319-325. https://doi.org/10.1177%2F0963721414541462

Dumford, A. D., & Miller, A. L. (2018). Online learning in higher education: Exploring advantages and disadvantages for engagement. *Journal of Computing in Higher Education, 30*(3), 452-465. http://doi.org/10.1007/s12528-018-9179-z

Dupere, V., Leventhal, T., Dion, E., Crosnoe, R., Archambault, I., & Janosz, M. (2015). Stressors and turning points in high school and dropout: A stress process, life course framework. *Review of Educational Research, 85*(4), 591-629. https://doi.org/10.3102/0034654314559845

Easton, J. Q., Johnson, E., & Sartain, L. (2017). *The predictive power of ninth-grade GPA.* Chicago, IL: University of Chicago Consortium on School Research. https://uchicago-dev.geekpak.com/sites/default/files/2018-10/Predictive%20Power%20of%20Ninth-Grade-Sept%202017-Consortium.pdf

Education Commission of the States. (2019). High school graduation requirements: What are the state's course requirements for high school graduation?. https://c0arw235.caspio.com/dp/b7f930000e16e10a822c47b3baa2

Ekwunife-Orakwue, K. C. V., & Teng, T. (2014). The impact of transactional distance dialogic interactions on student learning outcomes in online and blended environments. *Computer and Education, 78*, 414-427. https://doi.org/10.1016/j.compedu.2014.06.011

Elyakim, N., Reychav, I., Offir, B., & McHaney, R. (2019). Perceptions of transactional distance in blended learning using location-based mobile devices. *Journal of Educational Computing Research, 57(*1), 131-169. 0https://doi.org/10.1177/0735633117746169

Ersen, Z. B., Ezentas, R., & Altun, M. (2018). Evaluation of the teaching environment for improve the geometric habits of mind of tenth grade students. *European Journal of Education Studies, 4*(6), 47-64. http://dx.doi.org/10.5281/zenodo.1239849

Everatt, J., Fletcher, J., & Fickel, L. (2019). School leaders' perceptions on reading, writing and mathematics in innovative learning environments. *Education 3-13, 47*(8), 906-919. https://doi.org/10.1080/03004279.2018.1538256

Ezen-Can, A., & Boyer, K. E. (2015, June 26-29). *Choosing to interact: Exploring the relationship between learner personality, attitude, and tutorial dialogue participation* [Paper presentation]. International Conference on Educational Data Mining (EDM), Madrid, Spain. https://files.eric.ed.gov/fulltext/ED560560.pdf

Farah, R., Amarpreet, K., Zuraisa, R., Muzirah, M., Ayub, A., Nida, S., Rejeki, S., & Irwan, N. (2021). Algebraic lab: Pedagogical tool to teach and learn algebra through game. Review of International Geographical Education Online, 11(4), 951-962. https://rigeo.org/submit-a-menuscript/index.php/submission/article/view/491/941

Forte, G. J., Schwandt, D. R., Swayze, S., Butler, J., & Ashcraft, M. (2016). Distance education in the U.S.: A paradox. *Turkish Online Journal of Distance Education, 17*(3), 16-30. https://doi.org/10.17718/tojde.95102

Francis, J. E. (2018). Linking rubrics and academic performance: An engagement theory perspective. *Journal of University Teaching and Learning Practice*, *15*(1). http://ro.uow.edu.au/jutlp/vol15/iss1/3

Frazelle, S. (2016). *Successful strategies for providing online credit recovery in Montana.* Portland, OR: Education Northwest, Regional Educational Laboratory Northwest. https://files.eric.ed.gov/fulltext/ED566109.pdf

Freeman, J., & Simonsen, B. (2015). Examining the impact of policy and practice interventions on high school dropout and school completion: A systematic review of the literature. *Review of Educational Research, 85*(2), 205-248. https://doi.org/10:3102/0034654314554431

Fruehwirth, J. C., Biswas, S., & Perreira, K. M. (2021). The Covid-19 pandemic and mental health of first-year college students: Examining the effect of Covid-19 stressors using longitudinal data. *PLoS ONE, 16*(3), e0247999. https://doi.org/10.1371/journal.pone.0247999

Fryer, L. K., & Bovee, H. N. (2018). Staying motivated to e-learn: Person- and variable-centered perspectives on the longitudinal risks and support. *Computer and Education, 120*, 227-240. https://doi.org/10.1016/j.compedu.2018.01.006

Furr, R. M., & Heuckeroth, S. (2019). The "Quantifying Construct Validity" procedure: Its role, value, interpretations, and computation. Assessment, 26(4), 555-566. http://doi.org/10.1177/1073191118820638

Garcia-Perez, M. A. (2012). Statistical conclusion validity: Some common threats and simple remedies. *Frontiers in Psychology, 3*(325), 1-11. https://doi.org/10.3389/fpsyg.2012.00325

Garrison, R. (2000). Theoretical challenges for distance education in the 21st century: A shift from structural to transactional issues. *International Review of Research in Open and Distance Learning, 1*(1), 1-17. https://doi.org/10.19173/irrodl.v1i1.2

Garrison, R. D., & Cleveland-Innes, M. (2005). Facilitating cognitive presence in online learning: Interaction is not enough. *American Journal of Distance Education, 19*(3), 133-148. http://doi.org/10.1207s15389286ajde1903_2

Garthwait, A. (2014). Pilot program of online learning in three small high schools: Considerations of learning styles. *The Electronic Journal of e-Learning, 12(*4), 353-365. https://www.academic-publishing.org/index.php/ejel/article/view/1704/1667

Gates, P. (2020). Equity and access in mathematics education. In: Lerman S. (Ed.) *Encyclopedia of Mathematics Education*. Springer, Cham. http://doi.org/10.1007/978-3-030-15789-0_58

Gelman, A., & Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business and Economic Statistics, 37*(3), 447-456. https://doi.org/10.1080/07350015.2017.1366909

Georgia Credit Recovery Program. (2016). Credit Recovery FAQs. http://www.gacreditrecovery.org/FAQs.aspx

Georgia Department of Education, Georgia Virtual. (2011). Withdrawals and refunds.https://gavs.gavirtualschool.org/helpdeskweb/index.php?/Knowledgebase/Article/View/20/0/withdrawals-and-refunds

Georgia Department of Education. (2012). Four-year adjusted cohort high school graduation rate: Frequently asked questions. https://www.gadoe.org/External-Affairs-and-Policy/communications/Documents/Four-year%20Adjusted%20Cohort%20High%20School%20Graduation%20Rate%20FAQ.docx

Georgia Department of Education. (2014). Statewide passing score (State Board of Education Rule 160-4-2-.13). http://www.gadoe.org/External-Affairs-and-Policy/State-Board-of-Education/SBOE%20Rules/160-4-2-.13.pdf

Georgia Department of Education. (2020a). FY 2021 FTE data collection data element detail cycles 1 & 3. https://www.gadoe.org/Technology-Services/Data-Collections/Documents/FTE%20Resources/FY2021/FY2021%20FTE%20Data%20Element%20Detail.pdf

Georgia Department of Education. (2020b). Georgia Milestones Assessment System. https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Pages/Georgia-Milestones-Assessment-System.aspx

Georgia Department of Education. (2020c). End of Course Tests (EOCT). https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Pages/EOCT.aspx

Georgia Department of Education. (2020d). Georgia Performance Standards (GPS). https://www.georgiastandards.org/standards

Georgia Department of Education. (2020e). FY2021 Data Collection: Frequently asked questions. https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Special-Education-Services/Documents/SELDA/2020-2021/Aug%2020/FY2021%20Data%20Reporting%20FAQs.pdf

Georgia Department of Education. (2020f). Student enrollment by grade level (PK-12). [Data Table]. https://oraapp.doe.k12.ga.us/ows-bin/owa/fte_pack_enrollgrade.entry_form

Georgia Department of Education. (2020g). Enrollment by Ethnicity/Race and Gender. [Data Table]. https://oraapp.doe.k12.ga.us/ows-bin/owa/fte_pack_ethnicsex_pub.entry_form

Georgia Department of Education. (2020h). Student enrollment by disability. [Data Table]. https://oraapp.doe.k12.ga.us/ows-bin/owa/fte_pack_swd_enroll_pub.entry_form

Georgia Department of Education. (2022a). Understanding the Georgia Milestones achievement levels. https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Assessment/Pages/achievement_levels.aspx

Georgia Department of Education. (2022b). Georgia mathematics graduation requirement guidance. https://www.gadoe.org/Curriculum-Instruction-and-Assessment/Curriculum-and-Instruction/Documents/Mathematics/GA-HS-Graduation-Requirement-Mathematics-Guidance.pdf

Georgia Virtual Learning. (2014). The history of Georgia Virtual School. https://docs.google.com/document/d/1Y8742ve0yc5LgiTTxo5vJuG9-wgj5ckxK1Q7n-xoufo/edit#

Gewertz, C. (2020, November 5). Making a school reopening decision and taking the heat. *Education Week*. https://www.edweek.org/leadership/making-a-school-reopening-decision-and-taking-the-heat/2020/11

Giossos, Y., Koutsouba, M., Lionarakis, A., & Skavantzos, K. (2009). Reconsidering Moore's transactional distance theory. *European Journal of Open, Distance, and E-Learning, 2*. http://www.eurodl.org/materials/contrib/2009/Giossos_Koutsouba_Lionarakis_Skavantzos.htm

Glen, S. (2014, December 16). Internal validity: Definition and Examples. https://www.statisticshowto.datasciencecentral.com/internal-validity/#targetText=Internal%20validity%20is%20a%20way,you%20have%20in%20your%20experiment.&targetText=Therefore%20your%20internal%20validity%20would%20be%20very%20low.

Glen, S. (2015a, September 30). External validity and definition. https://www.statisticshowto.datasciencecentral.com/external-validity/

Glen, S. (2015b, September 30). Statistical conclusion validity. https://www.statisticshowto.datasciencecentral.com/statistical-conclusion-validity/

Goel, L., Zhang, P., & Templeton, M. (2012). Transactional distance revisited: Bridging face and empirical validity. *Computer in Human Behavior*, *28*, 1122-1129. https://doi.org/10.1016/j.chb.2012.01.020

Goodwin, R. N., Li, W., Broda, M., Johnson, H. L., & Schneider, B. (2016). Improving college enrollment of at-risk students at the school level. *Journal of Education for Students Placed at Risk*, *21*(3), 143-156. https://doi.org/10.1080/10824669.2016.1182027

Gopal, R., Singh, V., & Aggarwal, A. (2021). Impact of online classes on the satisfaction and performance of students during the pandemic period of COVID 19. *Education and Information Technologies*, *26*(1), 6923-6947. https://doi.org/10.1007/s10639-021-10523-1

Gopalan, M., Rosinger, K., & Ahn, J. B. (2020). Use of quasi-experimental research designs in education research: Growth, promise, and challenges. *Review of Research in Education*, *44*, 218-243. https://doi.org/10.3102/0091732X20903302

Gorsky, P., & Caspi, A. (2005). A critical analysis of transactional distance theory. *The Quarterly Review of Distance Education*, *6(*1), 1-11. https://www.openu.ac.il/personal_sites/download/avner-caspi/Gorsky&Caspi05.pdf

Governor's Office of Student Achievement. (2020). Downloadable data explained: Graduation and dropout rate. https://gosa.georgia.gov/sites/gosa.georgia.gov/files/OBIEE-Help/Graduation___Dropout_Rate.htm

Grace-Odeleye, B., & Santiago, J. (2019). A review of some diverse models of summer bridge programs for first-generation and at-risk college students. *Administrative Issues Journal: Connecting Education, Practice, and Research 9*(1), 35-47. https://doi.org/10.5929/9.1.2

Grad Inflation; Education. (2019, January 5). *The Economist, 430*(9124), 18(US). https://link.gale.com/apps/doc/A568150154/EAIM?u=minn4020&sid=ebsco&xid=333176b5

Graduation Alliance. (2017, March 6). The true cost of high school dropouts. https://www.graduationalliance.com/2017/03/06/the-true-cost-of-high-school-dropouts/

Grant, S. W., Hickey, G. L., & Head, S. J. (2019). Statistical primer: Multivariate regression considerations and pitfalls. European Journal of Cardi-Thoracic Surgery, 55(2), 179-185. https://doi.org/10.1093/ejcts/ezy403

Grigoryan, A. (2017). Audiovisual commentary as a way to reduce transactional distance and increase teaching presence in online writing instruction: Student perceptions and preferences. *Journal of Response to Writing, 3*(1), 83-128. https://scholarsarchive.byu.edu/journalrw/vol3/iss1/5/

Gubbels, J., van der Put, C., & Assink, M. (2019). Risk factors for school absenteeism and dropout: A Meta-analytic review. Journal of Youth and Adolescence, 48, pp. 1637-1667. https://link.springer.com/article/10.1007/s10964-019-01072-5

Guerda, O. (2020). An evaluation of construct and statistical conclusion validity in EFL research [Master's thesis, Larbi Ben M'hidi University - Oum El Bouagha]. http://bib.univ-oeb.dz:8080/jspui/bitstream/123456789/9548/1/thesis.pdf

Hageman, A. (2017). A review of the strengths and weaknesses of archival, behavioral, and qualitative research methods: Recognizing the potential benefits of triangulation. http://iranarze.ir/wp-content/uploads/2017/01/5994-English-IranArze.pdf

Hair, J., Anderson, R., Tatham, R., & Black, W. (1995). *Multivariate data analysis* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Halinski, R. S., & Feldt, L. S. (1970). The selection of variables in multiple regression analysis. *Journal of Educational Measurement, 7(*3), 151-158. https://doi.org/10.1111/j.1745-3984.1970.tb00709.x

Harrell, F., Lee, K. L., Matchar, D. B., & Reichert, T. A. (1985). Regression models for prognostic prediction: Advantages, problems and suggested solutions. *Cancer Treatment Reports, 69* (10), 1071-1077. PMID: 4042087

Hart, C. M. D., Berger, D., Jacob, B., Loeb, S., & Hill, M. (2019). Online learning, offline outcomes: Online course taking and high school student performance. *AERA Open, 5*(1). http://doi.org/10.1177/2332858419832852

Hassan, S., Waheed, H., Aljohani, N. R., Ali, M., Ventura, S., & Herrera, F. (2019). Virtual learning environment to predict withdrawal by leveraging deep learning. *International Journal of Intelligent Systems, 34*(8), 1935-1952. htts://doi.org/10.1002/int.22129

Heckman, J. J., Humphries, J. E., & Veramendi, G. (2018). The nonmarket benefits of education and ability. *Journal of Human Capital, 12*(2). https://www.journals.uchicago.edu/doi/abs/10.1086/697535

Heinrich, C. J., Darling-Aduana, J., Good, A., & Cheng, H. (2019). A look inside online educational settings in high school: Promise and pitfalls for improving educational opportunities and outcomes. *American Educational Research Journal, 56*(6), 2147-2188. https://doi.org/10.3102%2F0002831219838776

Heissel, J. (2016). The relative benefits of live versus online delivery: Evidence from virtual algebra I in North Carolina. *Economics of Education Review, 53*, pp 99-115. https://doi.org/10.1016/j.econedurev.2016.05.001

Hellerstein, J. M. (2008). Quantitative data cleaning for large databases. *United Nations Economic Commission for Europe*. https://dataresponsibly.github.io/courses/documents/Hellerstein_2008.pdf unece.pdf&ved=2ahUKEwiG2ZnUqb3mAhX_HiQlHesLDQEQFjAKegQlBhAB &usg=A0vVaw2SkNQSoWe9PdvHxW4e5N2w

Heppen, J. B., Sorenson, N., Allensworth, E., Walters, K., Rickles, J., Statchel Taylor, S., & Michelman, V. (2017). The struggle to pass algebra: Online vs f2f credit recovery for at-risk urban students. *Journal of Research on Educational Effectiveness, 10*(2), 272-296. http://dx.doi.org/10.1080/19345747.2016.1168500

Hillman, D. C. A., Willis, D. J., & Gunawardena, C. N. (1994). Learner-interface interaction in distance education: An extension of contemporary models for strategies and practitioners. *American Journal of Distance Education, 8*(2), pp. 30-42. https://doi.org/10.1080/08923649409526853

Horton, J. (2015). Identifying at-risk factors that affect college student success. *International Journal of Process Education, 7*(1), 1 – 19. https://pcrest.com/research/x2015%20Identifying%20At-Risk%20Factors%20That%20Affect%20College%20Student%20Success.pdf

Hoskins, B. J. (2012). Connections, engagement, and presence. *The Journal of Continuing Higher Education, 60*, 51-53. https://doi.org/10.1080/07377363.2012.650573

Huang, X., Chandra, A., DePaolo, C., Cribbs, J., & Simmons, L. (2015). Measuring transactional distance in web-based learning environments: An initial instrument development. *Open Learning: The Journal of Open, Distance and e-Learning, 30*(2), 106-126. https://doi.org/10.1080/02680513.2015.1065720

Huang, Z. (2018). Theoretical analysis of TPACK knowledge structure of mathematics teachers based on T-TPACK mode. *Educational Sciences: Theory and Practice, 18*(5), 2044-2053. https://doi.org/10.12738/estp.2018.5.103

Hughes, J., Zhou, C., & Petscher, Y. (2015). Comparing success for general and credit recovery courses online and f2f: Results for Florida high school courses (REL

2015-095). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. http://ies.ed.gov/ncee/edlabs.

Illuminate Education. (2019). aMath Score Interpretation Guide. https://fastbridge.illuminateed.com/hc/en-us/articles/1260802464050-aMath-Score-Interpretation-Guide

Illuminate Education. (2022a). Screening Basics. https://fastbridge.illuminateed.com/hc/en-us/articles/1260802463470-Screening-Basics

Illuminate Education. (2022b). aMath Overview. https://fastbridge.illuminateed.com/hc/en-us/articles/1260802450830-aMath-Overview

Illuminate Education. (2022c). Differential Item Function (DIF) Analysis: aReading and aMath. https://fastbridge.illuminateed.com/hc/en-us/articles/1260802304670-Differential-Item-Function-DIF-Analysis-aReading-and-aMath

Illuminate Education. (2022d). Assessment Overview. https://fastbridge.illuminateed.com/hc/en-us/articles/1260802411070-Assessment-Overview

Illuminate Education. (2022e). aMath. Retrieved January 31, 2022, from https://www.illuminateed.com/products/fastbridge/math-assessment/amath/

iNACOL. (2011). The online learning definitions project. Retrieved July 21, 2019, from https://www.inacol.org/wp-content/uploads/2015/02/iNACOL_DefinitionsProject.pdf

iNACOL. (2019). Our Work. Retrieved October 17, 2019, from https://www.inacol.org/our-work/

Irwin, S., Johnson, J., & Acker, B. (2014, September 19). *Meaningful, Engaged Learning.* Panel Discussion. Auburn University. Alabama. https://www.learner.org/workshops/socialstudies/pdf/session6/6.MeaningfulLearning.pdf

Ismail, A. O., Mahmood, A. K., & Abdelmaboud, A. (2018). Factors influencing academic performance of students in blended and traditional domains. *International Journal of Emerging Technologies in Learning (iJET)*, *13*(02), pp. 170–187. https://doi.org/10.3991/ijet.v13i02.8031

Jaster, R. W. (2017). Student and instructor perceptions of a flipped college algebra classroom. *International Journal of Teaching & Learning in Higher Education, 29*(1), 1-16. https://files.eric.ed.gov/fulltext/EJ1135836.pdf

Kaplon-Schilis, A., & Lyublinskaya, I. (2020). Analysis of relationship between five domains of TPACK framework: TK, PK, CK Math, CK Science, and TPACK of pre-service special education teachers. *Technology, Knowledge, and Learning, 25,* 25-43. http://doi.org/10.1007/s10758-019-09404-x

Kassandrinou, A., Angelaki, C., & Mavroidis, I. (2014). Transactional distance among open university students: How does it affect the learning process?. *European*

Journal of Open, Distance and E-Learning, *17*(1), 26-42. https://sciendo.com/downloadpdf/journals/eurodl/17/1/article-p26.pdf?pdfJsInlineViewToken=433040671&inlineView=true

Kearsley, G., & Lynch, W. (1996). Structural issues in distance education. *Journal of Education for Business*, *71*(4), 191. https://doi.org/10.1080/08832323.1996.10116783

Kearsley, G., & Schneiderman, B. (1998). Engagement theory: A framework for technology based teaching and learning. *Educational Technology, 38*(5), 20-23. https://www.jstor.org/stable/44428478

Kendall, J. M. (2003). Designing a research project: Randomized controlled trials and their principles. *Emerging Medicine Journal, 20*, 164-168. https://emj.bmj.com/content/emermed/20/2/164.full.pdf

Kim, C., Park, S., & Cozart, C. (2014). Affective and motivational factors of learning in online mathematics courses. *British Journal of Educational Technology, 45*(1), 171-185. https://doi.org/10.1111/j.1467-8535.2012.01382.x

Kim, D., Jung, E., Yoon, M., Chang, Y., Park, S., Kim, D., & Demir, F. (2021). Exploring the structural relationships between course design factors, learner commitment, self-directed learning, and intentions for further learning in a self-paced MOOC. *Computer and Education, 166*. https://doi.org/10.1016/j.compedu.2021.104171

Knowlton, D. S. (2000). A theoretical framework for the online classroom: A defense and delineation of a student-centered pedagogy. *New Directions for Teaching and Learning, 84*, 5-14. https://doi.org/10.1002/tl.841

Koehler, M. J., & Mishra, P. (2005). What happens when teachers design educational technology? The development of technological pedagogical content knowledge. *Journal for Educational Computing Research, 32*(2), pp. 131-152. https://doi.org/10.2190/0EW7-01WB-BKHL-QDYV

Kontkanen, S., Dillon, P., Valtonen, T., Eronen, L., Koskela, H., & Väisänen, P. (2017). Students' experiences of learning with iPads in upper secondary school – a base for proto-TPACK. *Education and Information Technologies, 22*(4), 1299-1326. https://doi.org/10.1007/s10639-016-9496-7

Kontorovich, I. (2018). Learning mathematics through online forums: A case of linear algebra. In E. Bergqvist, M. Osterholm, C. Granberg, & L. Sumpter (Eds.). *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education, 3*, 235-242. Umea, Sweden: PME. https://researchspace.auckland.ac.nz/bitstream/handle/2292/52732/Kontorovich_2018_Learning%20mathematics%20through%20online%20forums.pdf?sequence=2

Krieger, C. (2017). *Using Moore's transactional distance theory to examine selected online co-curricular educational opportunities in student affairs* (Unpublished doctoral dissertation). Virginia Polytechnic Institute and State University, Blacksburg, Virginia. http://hdl.handle.net/10919/80474

Kuhfeld, M., Gershoff, E., & Paschall, K. (2018). The development of racial/ethnic and socioeconomic achievement gaps during the school years. Journal of Applied Developmental Psychology, 57, 62-73. http://doi.org/10.1016/j.appdev.2018.07.001

Kumar, V., Alshazly, H., Idris, S. A., & Bourouis, S. (2021). Evaluating the impact of COVID-19 on society, environment, economy, and education. Sustainability, 13(24), 13642. https://doi.org/10.3390/su132413642

Laerd Dissertation. (2012). Construct validity. Statistical tutorials and software guides. http://dissertation.laerd.com/construct-validity.php

Laerd Statistics. (2015). Ordinal logistic regression using SPSS Statistics. Statistical tutorials and software guides. https://statistics.laerd.com

Laerd Statistics. (2017). Binomial logistic regression using SPSS Statistics. Statistical tutorials and software guides. https://statistics.laerd.com/

Laerd Statistics. (2018). Multinomial logistic regression using SPSS statistics. Statistical tutorials and software guides. https://statistics.laerd.com

Lara, J., Pelika, S., & Coons, A. (2017). Status of online learning programs in K-12: Implications for teachers. NEA Research Brief: NBI No. 135. https://www.nea.org/assets/docs/Online%20Learning%20Programs%20Research%20Brief%20NBI%20135%202017.pdf

Larkin, K., & Jamieson-Proctor, R. (2015). Using transactional distance theory to redesign an online mathematics education course for pre-service primary teachers.

Mathematics Teacher Education and Development, 17(1), 44-61. https://files.eric.ed.gov/fulltext/EJ1078418.pdf

Learning Theories and Student Engagement. (2014). *ASHE Higher Education Report, 40*(6), 15-36. http://doi-org.ezp.waldenlibrary.org/10.1002/aehe.20018

Leatherdale, S. (2019). Natural experiment methodology for research: A review of how different methods can support real-world research. International Journal of Social Research Methodology, 22(1), 19-35. https://doi.org/10.1080/13645579.2018.148849

Lesko, C. R., Buchanan, A. L., Westreich, D., Edwards, J. K., Hudgens, M. G., & Cole, S. R. (2017, July). Generalizing study results: A potential outcomes perspective. *Epidemiology, 28(*4), 553-561. https://doi.org/10.1097/EDE.0000000000000664

Li, Z. (2014). Power and sample size calculations for logistic regression tests for differential item functioning. *Journal of Educational Measurement, 51*(4), 441-462. https://doi.org/10.1111/jedm.12058

Libasin, Z., Azudin, A. R., Idris, N. A., Rahman, M. S. A., & Umar, N. (2021). Comparison of students' academic performance in mathematics course with synchronous and asynchronous online learning environments during COVID-19 crises. *International Journal of Academic Research in Progressive Education and Development*, *10*(2), 492-501. http://dx.doi.org/10.6007/IJARPED/v10-i2/10131

Lieberman, M. (2020, November 11). How hybrid learning is (and is not working) during COVID-19: 6 Case studies. *Education Week*.

https://www.edweek.org/leadership/how-hybrid-learning-is-and-is-not-working-during-covid-19-6-case-studies/2020/11

Limtrairut, P., & Marshall, S. (2020, October 13-16). A new design guideline for mobile learning application: Transactional distance perspective. 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE), pp. 610-614. https://doi.org/10.1109/GCCE50665.2020.9291976

Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage.

Los Angeles Mission College, Office of Institutional Research and Planning. (2011). A brief guide to data definition and calculation for enrollment, instructional efficiency, and student achievement. http://www.lamission.edu/irp/docs/Data%20Definition.pdf

Los Angeles Southwest College. (2019). Equity Plan and Student Success Indicators. https://www.lasc.edu/equity/Equity-Plan-and-Student-Success-Indicators

Lu, O. H. T., Huang, A. Y. Q., & Yang, S. J. H. (2021). Impact of teachers' grading policy on the identification of at-risk students in learning analytics. Computers and Education, 163, 1-13. https://doi.org/10.1016/j.compedu.2020.104109

Lu, Y. (2011). Using a virtual classroom to teach online mathematics. https://eric.ed.gov/?id=ED519767

Lucky, A., Branham, M., & Atchison, R. (2019). Collection-based education by distance and f2f: Learning outcomes and academic dishonesty. *Journal of Science*

Education and Technology, *28*(1), 414–428. https://doi.org/10.1007/s10956-019-9770-8

Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, *1*(3), 86-92. https://psycnet.apa.org/doi/10.1027/1614-2241.1.3.86

Macaskill, A., & Taylor, E. (2010). The development of a brief measure of learner autonomy in university students. *Studies in Higher Education, 35*(3), 351-359. https://doi.org/10.1080/03075070903502703

MacLeod, K., Swart, W., & Paul, R. (2019). Continual improvement of online and blended teaching using relative proximity theory. *Decision Sciences Journal of Innovative Education, 17*(1), 53-75. https://eds.p.ebscohost.com/eds/pdfviewer/pdfviewer?vid=1&sid=3a5038e4-2cb8-409b-962b-ba04fb103955%40redis

Malkus, N. (2018). Second chance to second track? Credit recovery participation in US high schools. *American Enterprise Institute*. https://www.aei.org/publication/second-chance-or-second-track-credit-recovery-participation-in-us-high-schools/

Mamedova, S., & Pawlowski, E. (2018). *A description of U.S. adults who are not digitally literate* (NCES Brief 2018-161). https://nces.ed.gov/pubs2018/2018161.pdf

Marange, T., & Adendorff, S. S. (2021). The contribution of online mathematics games to algebra understanding in Grade 8. *Pythagoras, 42*(1), 8. https://journals.co.za/doi/abs/10.4102/pythagoras.v42i1.586

Marcus, J. (2016, August 18). A solution as obvious as it is rare: Making high school graduates ready for college. The Hechinger Report. https://hechingerreport.org/solution-obvious-rare-making-high-school-graduates-ready-college/

Martin, F., & Bolliger, D. U. (2018). Engagement matters: Student perceptions on the importance of engagement strategies in the online learning environment. *Online Learning 22*(1), 205- 222. https://doi.org/10.24059/olj.v22i1.1092

Martin, F., & Kumar, S. (2018). Frameworks for Assessing and Evaluating e-Learning Courses and Programs. In: Peña, A., Lowell, V., Harris, B. (eds) Leading and Managing e-Learning. Educational Communications and Technology: Issues and Innovations. Springer, Cham. http://doi.org/10.1007/978-3-319-61780-0_19

Mascha, E., & Vetter, T. R. (2018). Significance, errors, power, and sample size: The blocking and tackling of statistics. Anethesia & Analgesia, 126(2), 691-698. https://doi.org/10.1213/ANE.0000000000002741

Matherson, L. H., Wilson, E. K., & Wright, V. H. (2014). Need TPACK? Embrace sustained professional development. *The Delta Kappa Gamma Bulletin*, 45- 52.

Mbwesa, J. K. (2014). Transactional distance as a predictor of perceived learner satisfaction in distance learning courses: A case study of Bachelor of Education Arts Program, University of Nairobi, Kenya. *Journal of Education and Training Studies, 2*(2), 176-188. https://files.eric.ed.gov/fulltext/EJ1055516.pdf

McGee, S., Greenberg, R. I., Dettori, L., Rasmussen, A. M., McGee-Tekula, R., Duck, J., & Wheeler, E. (2018). An examination of factors correlating with course failure in a high school computer science course. *The Learning Partnership, Tech. Rep, 5*. https://www.jointhepartnership.net/wp-content/uploads/2020/07/Factors-Correlated-with-Course-Failure.pdf

McKenzie, S. C., McGee, J. B., Reid, C. A., & Morris, S. R. (2022). Arkansas high school freshman course failures [Policy brief]. University of Arkansas, Fayetteville. *Office for Education Policy*, *19*(3). https://scholarworks.uark.edu/oepbrief/162/

McKinley, J., & Rose, H. (Eds.). (2019). The Routledge Handbook of Research Methods in Applied Linguistics (1st ed.). Routledge.

McMillion, T., & King, C. (2017). Communication and security issues in online education: Student self-disclosure in course introductions. *Journal of Interactive Online Learning, 15(*1). https://eds.p.ebscohost.com/eds/pdfviewer/pdfviewer?vid=1&sid=127f016c-9ae5-45b4-bd77-fa648607a632%40redis

Meissel, K., Meyer, F., Yao, E. S., & Rubie-Davies, C. M. (2017). Subjectivity of teacher judgements: Exploring students characteristics that influence teacher judgements of student ability. *Teaching and Teacher Education, 65*, 48-60. https://doi.org/10.1016/j.tate.2017.02.021

Menchaca, M. P., & Bekele, T. A. (2008). Learner and instructor identified success factors in distance education. *Distance Education, 29*(3), 231-252. https://doi.org/10.1080/01587910802395771

Michael, R. S. (n.d.). Threats to internal and external validity. Indiana University. http://www.indiana.edu/~educy520/sec5982/week_9/520in_ex_validity.pdf

Miller, D. E., & Kunce, J. T. (1973). Prediction and statistical overkill revisited. *Measurement and Evaluation in Guidance, 6*(3), 157-163. https://doi.org/10.1080/00256307.1973.12022590

Miron, G., & Gulosino, C. (2016). Virtual schools report 2016: Directory and performance review. Boulder, CO: National Education Policy Center. http://nepc.colorado.edu/publication/virtual-schools-annual-2015

Mollenkopf, D., Vu, P., Crow, S., & Black, C. (2017). Does online learning deliver? A comparison of student teacher outcomes from candidates in f2f and online program pathways. Online Journal of Distance Learning Administration, 20(1), n1. https://ojdla.com/archive/spring201/mollenkopf_vu_crow_black201.pdf

Molnar, A., Miron, G., Elgeberi, N., Barbour, M. K., Huerta, L., Shafer, S. R., & Rice, J. K. (2019). *Virtual Schools in the U.S. 2019*. Boulder, CO: National Education Policy Center. http://nepc.colorado.edu/sites/default/files/publications/Virtual%20Schools%202019.pdf

Moore, M. (1997). Theory of transactional distance. In D. Keegan (Ed.). *Theoretical Principles of Distance Education* (pp. 22-38). Routledge.

Moore, M. G. (1972). Learner autonomy: the second dimension of independent learning. *Convergence*, *2*, 76-88.

Moore, M. G. (1989). Three types of interaction. *American Journal of Distance Education, 3*(2), 1-7. https://doi.org/10.1080/08923648909526659

Moore, M. G. (1993). Theory of transactional distance. *Theoretical Principals of Distance Education, 22.*

Muhonen, H., Pakarinen, E., Poikkeus, A., Lerkkanen, M., & Rasku-Puttonen, H. (2018). Quality of educational dialogue and association with students' academic performance. *Learning and Instruction, 55*(1), 67-79. https://doi.org/10.1016/j.learninstruc.2017.09.007

Murphy, E., & Rodriguez-Manzanares, M. A. (2008). Revisiting transactional distance theory in a context of web-based high-school distance education. *Journal of Distance Education, 22(*2), 1-14. https://www.ijede.ca/index.php/jde/article/view/38/549

National Center for Educational Statistics. (2015). *Percentage of public school districts with students enrolled in technology-based distance education courses and number of enrollments in such courses, by instructional level and district characteristics: 2002-03, 2004-05, and 2009-10* (Table 218.20). U.S. Department of Education. Washington, DC. https://nces.ed.gov/programs/digest/d14/tables/dt14_218.20.asp

National Center for Education Statistics. (2020). *Number of virtual schools, total state enrollment, total virtual school enrollment as a percentage of state total enrollment: School year 2019-2020* (Table 3). U.S. Department of Education.

Washington, DC. https://nces.ed.gov/ccd/tables/201920_Virtual_Schools_tables_3.asp

National Center for Education Statistics. (2021a). *The Condition of Education 2021* (NCES 2021-144), Status Dropout Rates. https://nces.ed.gov/fastfacts/display.asp?id=16

National Center for Education Statistics. (2021b). Population 16 to 24 year old and number of 16- to 24-year-old high school dropouts (status dropouts), by sex and race/ethnicity: Selected years 1970 through 2019. [Data table 219.71]. https://nces.ed.gov/programs/digest/d20/tables/dt20_219.71.asp

National Forum on Education Statistics, Race/Ethnicity Data Implementation Task Force. (2008). Managing an Identity Crisis: Forum Guide to Implementing New Federal Race and Ethnicity Categories (NFES 2008- 802). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. https://nces.ed.gov/pubs2008/2008802.pdf

Newkirk, H. E., Schwager, P. H., & Eakins, L. (2013). An empirical investigation of perceptions and achievements of students in online and f2f courses. *American Journal of Educational Studies, 6*(1), 33-45. http://www.paulschwager.com/pubs/2013AJES.pdf

Niemi, H., & Kousa, P. (2020). A case study of students' and teachers' perceptions in a Finnish high school during the COVID pandemic. *International Journal of Technology in Education and Science, 4*(4), pp. 352-369. http://doi.org/10.46328/ijtes.v4i4.167

Niu, L. (2020). A review of the application of logistic regression in educational research: Common issues, implications, and suggestions. *Educational Review, 72*(1), 41-67. https://doi.org/10.1080/00131911.2018.1483892

Noble, K., Pelika, S., & Coons, A. (2017). Online credit recovery programs: NEA research brief: NBI No. 2. https://www.nea.org/assets/docs/Online%20Credit%20Recovery%20Programs.%20Research%20Brief.%20NBI%202%202017.pdf

Office of Planning, Evaluation and Policy Development. (2017). National survey on high school strategies designed to help at-risk students graduate. https://www2.ed.gov/about/offices/list/opepd/ppss/reports-high-school.html

Oliver, K., & Kellogg, S. (2015). Credit recovery in virtual school: Affordances of online learning for the at-risk student. *Journal of Online Learning Research, 1*(2), 191-218. https://www.learntechlib.org/primary/p/149111/

Oliver, K., Osborne, J., Patel, R., & Kleiman, G. (2009). Issues Surrounding the Deployment of a New Statewide Virtual Public School. *Quarterly Review of Distance Education, 10*(1), 37-49. https://www.researchgate.net/profile/Jason-Osborne/publication/265963847_Issues_surrounding_the_deployment_of_a_new_statewide_virtual_public_school/links/55f70cc908aeba1d9eef8683/Issues-surrounding-the-deployment-of-a-new-statewide-virtual-public-school.pdf

Olvera Astivia, O. L., Gadermann, A., & Guhn, M. (2019). The relationship between statistical power and predictor distribution in multilevel logistic regression: A simulation-based approach. *BMC Medical Research Methodology*, *19*(97). https://doi.org/10.1186/s12874-019-0742-8

Organization for Economic Cooperation and Development. (2018). Skills for jobs. https://www.oecd.org/els/emp/Skills-for-jobs-brochure-2018.pdf

Organization for Economic Cooperation and Development. (2021). OECD employment outlook 2021: Navigating the COVID-19 crises and recovery. https://doi.org/10.1787/5a700c4b-en

Organization for Economic Cooperation and Development. (2022). Employment by education level (indicator). https://doi.org/10.1787/26f676c7-en

Ostashevsky, L. (2016, February 18). Many who pass state high school graduation tests show up to college unprepared. The Hechinger Report. https://hechingerreport.org/many-who-pass-state-high-school-graduation-tests-show-up-to-college-unprepared/

Oviatt, D. R., Graham, C. R., Borup, J., & Davies, R. S. (2016). Online student perceptions of the need for proximate community of engagement at an independent study program. *Journal of Online Learning Research*, *2*(4), 333-365. https://files.eric.ed.gov/fulltext/EJ1148598.pdf

Owan, V. J., & Asuquo, M. E. (2021). Assessment of Socio-Demographic Factors and Students’ Satisfaction with the Study of ICT in Secondary Schools. Pedagogical Research, 6(3), em0101. https://doi.org/10.29333/pr/11087

Papadopoulos, I., & Dagdilelis, V. (2006). The theory of transactional distance as a framework for the analysis of computer-aided teaching of geometry. *International Journal for Technology in Mathematics Education, 13*(4), 175-182.

Paschall, K. W., Gershoff, E. T., & Kuhfeld, M. (2018). A two decade examination of historical race/ethnicity disparities in academic achievement by poverty status. *Journal of Youth and Adolescence, 47*, 1164-1177. https://doi.org/10.1007/s10964-017-0800-7

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach.* Hillside, NJ: Lawrence Erlbaum.

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology, 49*(12), 1373-1379. https://doi.org/10.4324/9780203726389

Picciano, A. G., Seaman, J., & Day, S. L. (2015). Online learning in Illinois high schools: The voices of principals!. In T.L. Heafner, R. Hartshorne, & T. Petty (Eds.) Exploring the Effectiveness of Online Education in K-12 Environments (pp. 1-18). Hershey, PA: IGI Global. https://doi.org/10.4018/978-1-4666-6383-1

Powell, A., Roberts, V., & Patrick, S. (2015). Using online learning for credit recovery: Getting back on track to graduation. Vienna, VA: International Association of K-12 Online Learning (iNACOL). https://www.inacol.org/wp-content/uploads/2015/09/iNACOL_UsingOnlineLearningForCreditRecovery.pdf

Protopsaltis, S., & Baum, S. (2019). Does online education live up to its promise? A look at the evidence and implications for federal policy. http://mason.gmu.edu/~sprotops/OnlineEd.pdf

Purdue University. (2017). Copyright exceptions: F2f instruction. https://www.lib.purdue.edu/uco/ForInstructors/face_to_face.html

Quong, J., Snider, S. L., & Early, J. (2018). Reducing transactional distance in online and blended courses through the use of a closed social media platform. *Journal of Educational Technology*, *47*(1), 79-100. https://doi.org/10.1177%2F0047239518766654

Repetto, J. B. (2018). Research on at-risk learners in K-12 online learning. In K. Kennedy & R. E. Ferdig (Eds.), *Handbook of research on K-12 online and blended learning* (2nd ed., pp. 163-180). Carnegie Mellon University: ETC Press.

The Research and Planning Group for California Community Colleges. (2011). Operational Definitions. https://rpgroup.org/Institutional-Research-Operational-Definitions

Rickaby, C. (2021). Algebra I or traditional math in eighth grade?: The effects on school district PSAT and SAT performance scores (Publication No. 28544537) [Doctoral dissertation, Southern Connecticut State University]. ProQuest Dissertations Publishing. https://www.proquest.com/openview/72bd4a92c10b52c042dd915e78a191fc/1?pq-origsite=gscholar&cbl=18750&diss=y

Rickles, J., Heppen, J. B., Allensworth, E., Sorenson, N., & Walters, K. (2018). Online credit recovery and the path to on-time high school graduation. *Educational Researcher, 47*(8), 481-491. https://doi.org/10.3102/0013189x18788054

Rosenzweig, E. Q., Hulleman, C. S., Barron, K. E., Kosovich, J. J., Priniski, S. J., & Wigfield, A. (2019). Promises and pitfalls of adapting utility value interventions for online math courses. *The Journal of Experimental Education, 87*(2), https://doi.org/10.1080/00220973.2018.1496059

Saba, F., & Shearer, R. L. (1994). Verifying key theoretical concepts in a dynamic model of distance education. *The American Journal of Distance Education, 8*(1), 36-59.

Saba, F., & Shearer, R. L. (2018). *Transactional distance and adaptive learning: Planning for the future of higher education.* Routledge. https://doi.org/10.4324/9780203731819

Salkind, N. J. (Ed.). (2010). Hawthorne effect. *Encyclopedia of Research Design.* https://methods.sagepub.com/reference/encyc-of-research-design/n174.xml

Santos, K. C. P., & Barrios, E. B. (2017). Improving predictive accuracy of logistic regression model using ranked set samples. *Communications in Statistics – Simulation and Computation, 46*(1), 78-90. http://doi.org/10.1080/03610918.2014.955113

Sarikas, C. (2015, September 13). Should you drop a class? Does it look bad? Why? [Web log]. https://blog.prepscholar.com/dropping-a-class-does-it-look-bad

Satsangi, R., Raines, A. R., & Fraze, K. (2021). Virtual manipulatives for teaching algebra: A research-to-practice guide for secondary students with a learning disability. *Learning Disabilities: A Multidisciplinary Journal, 26*(1). https://doi.org/10.18666/LDMJ-2021-V26-I1-10349

Schwartz, S. (2018, February 7). Educator development key to statewide personalized learning, iNACOL says. http://blogs.edweek.org/edweek/DigitalEducation/2018/02/personalized_learning_professional_development.html

Serdyukov, P. (2015). Does online education need a special pedagogy? *Journal of Computing and Information Technology, 23*(1), 61-74. https://doi.org/10.2498/cit.1002511

Shin, N. (2002). Beyond interaction: The relational construct of transactional presence. *Open Learning, 17*(2), 122-137. https://doi.org/10.1080/02680510220146887

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14. https://doi.org/10.3102%2F0013189X015002004

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57(*1), 1-21. https://doi.org/10.17763/haer.57.1.j463w79r56455411

Singh, V., & Thurman, A. (2019). How many ways can we define online learning? A systematic literature review of definitions of online learning (1988-2018). *American Journal of Distance Education, 33*(4), 289-306. https://doi.org/10.1080/08923647.2019.1663082

Smith Jaggars, S., Edgecombe, N., & West Stacey, G. W. (2013). *What we know about online course outcomes.* Community College Research Center, Columbia University. https://files.eric.ed.gov/fulltext/ED542143.pdf

Sorubakhsh-Castillo, M. (2018). A case study: How freshman teachers perceive grade inflation at a southeast high school. [Dissertation, Concordia University]. Available from ProQuest One Academic (20190668583). https://www.proquest.com/dissertations-theses/case-study-how-freshman-teachers-perceive-grade/docview/2190668583/se-2

Stein, D. S., Wanstreet, C. E., Calvin, J., Overtoom, C., & Wheaton, J. E. (2005). Bridging the transactional distance gap in online learning environments. *American Journal of Distance Education, 19*(2), 105-118. https://doi.org/10.1207/s15389286ajde1902_4

Stevens, D., Frazelle, S., Bisht, B., & Hamilton, R. (2016). *Online credit recovery: Enrollment and passing patterns in Montana Digital Academy courses* (REL 2016-139) Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest. https://ies.ed.gov/ncee/rel/regions/northwest/pdf/REL_2016139.pdf

Straub, M. L. (2021). *The relationships between scores on FastBridge aReading and aMath and scores on the Kansas English Language Arts and Mathematics assessments in grades 3-8.* [Unpublished Doctoral Dissertation, Baker

University]. https://www.bakeru.edu/images/pdf/SOE/EdD_Theses/Straub_Misty.pdf

Sutton, A., Langenkamp, A. G., Muller, C., & Schiller, K. S. (2018). Who gets ahead and who falls behind during the transition to high school? Academic performance at the intersection of race/ethnicity and gender. Social Problems, 65(2), 154-173. http://doi.org/10.1093/socpro/spx044

Swart, W., & MacLeod, K. (2021). Evaluating learning space designs for flipped and collaborative learning: A transactional distance approach. *Education Sciences, 11*(292). https://doi.org/10.3390/educsci11060292

Thai, N. T. T., De Wever, B., & Valcke, M. (2020). F2f, blended, flipped, or online learning environment? Impact on learning performance and student cognitions. *Journal of Computer Assisted Learning, 36*(3). https://doi.org/10.1111/jcal.12423

Timberlake, A. M. (2011). *Sample size in ordinal logistic hierarchical linear modeling*. (Dissertation, Georgia State University). http://scholarworks.gsu.edu/eps_diss/72

Tirronen, V., Lappalainen, V., Isomöttönen, V., Lakanen, A., Taipalus, T., Nieminen, P., & Ogbechie, A. (2020, October 21-24). *Incorporating teacher-student dialogue into digital course material: Usage patterns and first experiences.* [Paper presentation]. *2020 IEEE Frontiers in Education Conference (FIE)*, 1-5. https://doi.org/10.1109/FIE44824.2020.9274123.

Toker, Z., & Bektaş, D. (2021). Reflections on online mathematics teaching. *Necatibey Faculty Education Electronic Journal of Science & Mathematics Education, 15*(2), 233-253. https://doi.org/10.17522/balikesirnef.964470

Tromski-Klingshirn, D., & Miura, Y. (2017). School counselor's role in dropout prevention and credit recovery [Special Edition]. *Journal of School Counseling, 15*(4), 1-17. https://files.eric.ed.gov/fulltext/EJ1144756.pdf

Umbara, U., Susilana, R., & Puadi, E. F. W. (2021). Algebra dominoes game: Re-designing mathematics learning during the COVID-19 pandemic. *International Journal of Instruction, 14*(4). https://files.eric.ed.gov/fulltext/EJ1319616.pdf

United Nations International Children's Emergency Fund. (2020). Education and COVID-19. https://data.unicef.org/topic/education/covid-19/

University of Northern British Columbia. (2012). How are enrollment statistics calculated?. https://www.unbc.ca/sites/default/files/assets/institutional_research/how_are_enrolment_statistics_calculated_2012_jul.pdf

U.S. Department of Education (ED). (2007). At-risk student intervention implementation guide: A comprehensive resource for identifying programs to help decrease South Carolina's school dropout population. https://www2.ed.gov/rschstat/eval/sectech/pathways2careers/AtRiskStudentGuide.pdf

Usta, N., & Mirasyedioğlu, S. (2022). The effect of distance education on the problem-solving success of university students taking algebra and students' opinions on the practice. *Journal of Education in Black Sea Region, 7*(2), 141-169. https://jebs.ibsu.edu.ge/jms/index.php/jebs/article/view/269/278

Valverde-Berrocoso, J., Garrido-Arroyo, M. C., Burgos-Videla, C., & Morales-Cevallos, M. B. (2020). Trends in educational research about e-learning: A systematic literature review (2009-2018). *Sustainability, 12(*12). https://doi.org/10.3390/su12125153

van Smeden, M., de Groot, J. A. H., Moons, K. G. M., Collins, G. S., Altman, D. G., Eijkemans, M. J. C., & Reitsma, J. B. (2016). No rational for 1 variable per 10 events criterion for binary logistic regression analysis. *BMC Medical Research Methodology, 16*(163), 1-12. https://doi.org/10.1186/s12874-016-0267-3

van Smeden, M., Moons, K. G. M., de Groot, J. A. H., Collins, G. S., Altman, D. G., Eijkemans, M. J. C., & Reitsma, J. B. (2019). Sample size for binary logistic models: Beyond events per variable criteria. Statistical Methods in Medical Research, 28(8), 2455-2474. https://doi.org/10.1177/0962280218784726

Velasquez, A., Graham, C. R., & West, R. E. (2013). An investigation of practices and tools that enabled technology-mediated caring in an online high school. *The International Review of Research in Open and Distance Learning*, *14*(5), 277-299. https://files.eric.ed.gov/fulltext/EJ1017544.pdf

Viano, S. L. (2018). At-risk high school students recovering course credits online: What we know and need to know. *American Journal of Distance Education, 32*(1), 16-26. https://doi.org/10.1080/08923647.2018.1412554

Vickrey, T., Golick, D., & Stains, M. (2018). Educational technologies and instructional practices in agricultural sciences: Leveraging the technological pedagogical content knowledge (TPACK) framework to critically review the literature. *North*

American Colleges and Teachers of Agriculture (NACTA), 62(1), 65-76. https://nactateachers.org/index.php/vol-62-1-mar-2018/2700-educational-technologies-and-instructional-practices-in-agricultural-sciences-leveraging-the-technological-pedagogical-content-knowledge-tpack-framework-to-critically-review-the-literature

Villares, E., & Brigman, G. (2019). College/Career success skills: Helping students experience postsecondary success. Professional School Counseling, 22(1b), 1-8. https://doi.org/10.1177%2F2156759X19834444

Wakil, K., Abdulfaraj, A., Sadula, A., Tofiq, D., & Nawzad, L. (2019). Performance of distance learning compared with f2f learning. *Journal of Educational Science and Technology, 5*(1). http://doi.org/10.26858/est.v5i1.7952

Walden University, Center for Research Quality (CRQ). (2019a). General tips for avoiding ethical problems in doctoral research. https://academicguides.waldenu.edu/researchcenter/orec/frequently-asked-questions/international#s-lg-box-21161416

Walden University, Center for Research Quality (CRQ). (2019b). Research Ethics and Compliance: FAQs educational settings. https://academicguides.waldenu.edu/researchcenter/orec/frequently-asked-questions/educational-settings

Waschull, S. B. (2005). Predicting success in online psychology courses: Self-discipline and motivation. *Teaching of Psychology, 32*(3), 190-192. https://doi.org/10.1207%2Fs15328023top3203_11

Weidlich, J., & Bastiaens, T. J. (2018). Technology matters – The impact of transactional distance on satisfaction in online distance learning. *The International Review of Research in Open and Distributed Learning*, *19*(3), 222-242. https://doi.org/10.19173/irrodl.v19i3.3417

Weiss, D. M., & Belland, B. R. (2018). PBL group autonomy in high school environmental science class. *Technology, Knowledge, and Learning*, *23*, 83-107. http://doi.org/10.1007/s10758-016-9297-5

Wengrowicz, N., & Offir, B. (2013). Teachers' perceptions of transactional distance in different teaching environments. *The American Journal of Distance Education, 27*, 111-121. https://doi.org/10.1080/08923647.2013.773701

What Works Clearinghouse. (2017a). Procedures Handbook Version 4.0. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_handbook_v4.pdf

What Works Clearinghouse. (2017b). *WWC review of the report: The struggle to pass Algebra: Online vs. f2f credit recovery for at-risk urban students*. https://ies.ed.gov/ncee/wwc/Docs/SingleStudyReviews/wwc_heppen_021417.pdf

Wheatley, D. M. (2016). *Virtual high school and instructional design strategies to reduce transactional distance and increase student engagement: A Delphi study* (Doctoral dissertation, Capella University). https://pqdtopen.proquest.com/doc/1846530525.html?FMT=AI

White, H., & Sabarwal, S. (2014). Quasi-experimental design and methods. *Methodological Briefs: Impact Evaluation 8,* UNICEF Office of Research, Florence. https://www.researchgate.net/profile/Sharanbasappa-Durg/post/Can_a_quasi-experimental_design_be_retrospective/attachment/59d63c4179197b80779993a1/AS%3A414509891112960%401475838267060/download/Brief_8_quasi-experimental+design_eng.pdf

Yates, A., Starkey, L., Egerton, B., & Flueggen, F. (2021). High school students' experience of online learning during COVID-19: the influence of technology and pedagogy. *Technology, Pedagogy, and Education*, *30*(1), 59-73. https://doi.org/10.1080/1475939X.2020.1854337

Yilmaz, F. G. K. (2017). Social presence and transactional distance as an antecedent to knowledge sharing in virtual learning communities. *Journal of Educational Computing*, *55*(6), 844-864. https://doi.org/10.1177/0735633116688319

Yilmaz, R., & Keser, H. (2017). The impact of interactive environment and metacognitive support on academic achievement and transactional distance in online learning. *Journal of Educational Computing*, *55*(1), 95-122. https://doi.org/10.1177/0735633116656453

Zb, A., Novalian, D., Ananda, R., Habibi, M., & Sulman, F. (2021). Distance learning with STEAM approaches: Is the effect on the cognitive domain? Journal of Educational Studies, 6(2), 128-138. http://dx.doi.org/10.30983/educative.v6i2.4977

Zhang, A. (2003). Transactional distance in web-based college learning environments: Towards measurement and theory construction. [Ph.D. Dissertation]. Virginia

Commonwealth University, Richmond, VA, USA. https://www.proquest.com/docview/305222372?pq-origsite=gscholar&fromopenview=true

Zimmerman, T. D. (2012). Exploring learner to content interaction as a success factor in online courses. *International Review of Research in Open and Distance Learning, 13*(4), 152-165. https://doi.org/10.19173/irrodl.v13i4.1302

Appendix

Certificate of Completion

The National Institutes of Health (NIH) Office of Extramural Research certifies that **Andrea Rohde** successfully completed the NIH Web-based training course “Protecting Human Research Participants”.

Date of completion: 12/23/2012

Certification Number: 1067437