6-2018

Addressing Ethnic and Gender Differences in Advanced Placement Calculus (AB) Exam Scores

Clint S. Prong
Concordia University - Portland

Concordia University–Portland

College of Education

Doctorate of Education Program

WE, THE UNDERSIGNED MEMBERS OF THE DISSERTATION COMMITTEE

CERTIFY THAT WE HAVE READ AND APPROVE THE DISSERTATION OF

Clint Stephen Prong

CANDIDATE FOR THE DEGREE OF DOCTOR OF EDUCATION

Neil Mathur, Ph.D., Faculty Chair Dissertation Committee

Jessica deValentino, Ed.D., Content Specialist

Nesa Sasser, Ed.D., Content Reader

Addressing Ethnic and Gender Differences in Advanced Placement Calculus (AB) Exam Scores

Clint Stephen Prong

Concordia University–Portland

College of Education

Dissertation submitted to the Faculty of the College of Education

in partial fulfillment of the requirements for the degree of

Doctor of Education in

Teacher Leadership

Neil Mathur Ph.D., Faculty Chair Dissertation Committee

Jessica deValentino, Ed.D., Content Specialist

Nesa Sasser, Ed.D., Content Reader

Concordia University–Portland

2018

Abstract

The purpose of this quantitative study was to determine was to test whether increasing the length of the instructional period from one to two hours might increase student achievement and close achievement gaps. This study included two components: an achievement gap analysis and a quasi-experiment. The purpose of the achievement gap analysis was to measure the existence of achievement gaps on the Calculus AB exam in Michigan. The achievement gap analysis included comparisons of Calculus AB exam scores in Michigan to measure gender differences, ethnic differences, and Hispanic gender differences. The purpose of the quasi-experiment was to determine whether increasing the length of the instructional period might close achievement gaps. In the quasi-experiment, the comparison group took Calculus for one hour daily, while the experiential group took Calculus for two hours daily. Both the comparison group and the experiential group took the Calculus AB exam at the end of their year of instruction. The results of the achievement gap analysis confirmed that Michigan has both gender and ethnic achievement gaps that showed up in every year of analyzed data, while a Hispanic achievement gap existed only in the most recent year of data. The results of the quasi-experimental component indicated that increasing the amount of Calculus instruction from one hour to two hours resulted in a statistically significant difference in all four comparisons of the comparison group to the experiential group. The results of this study indicated that schools should consider increasing the length of the period of Calculus instruction to address achievement gaps.

Keywords: Advanced Placement, gender differences, ethnic differences, calculus achievement

Dedication

I dedicate this dissertation to my loving wife, Heather. Without your companionship, love, and encouragement, the completion of this work would not have been possible. Thank you for your constant support throughout the process of balancing work and the pursuit of a doctoral degree.

Acknowledgements

No individual ever completes a dissertation without aid and support. First, I could never have persisted through this doctoral journey except for the saving grace of my Lord Jesus Christ. I pray that my efforts, both in pursuit of a doctorate and in my personal interactions, will always bring glory to Him!

Additionally, I owe a debt of gratitude to the following Concordia University–Portland personnel. Dr. Mathur, thank you for serving as my dissertation chair and thank you for your unflagging support and encouragement throughout the process of developing, writing, and editing my dissertation. Dr. deValentino, thank you for your willingness to serve as my content specialist and your guidance in the design of the study in this dissertation. Dr. Sasser, thank you for your timely feedback and prompt communication. Dr. Bullis, thank you for taking the time to address my questions and concerns throughout the dissertation process. Dr. Owusu-Ansah, thank you for sharing your advice and expertise to expedite the editions of my dissertation.

Closer to home, I also owe a debt of gratitude to the following people. Family and friends, thank you for your continuous support through the entirety of the doctoral journey. Former Calculus students, thank you for your enthusiasm and support of my pursuit of my doctorate.

Table of Contents

List of Tables

Chapter 1: Introduction

The importance of an education can hardly be overstated. Addressing that point, President Obama (2016) instructed graduates of one high school that, "…a high school education these days is not enough. By 2020, two out of three job openings require some form of higher education" (para. 19). However, the attainment of an education does not only affect employability, but also the compensation that accompanies that employment (Autor, 2014; Acemoglu & Autor, 2012; Obama, 2016). Discrepancies in levels of educational attainment continue to widen the income gap (Acemoglu & Autor, 2012; Autor, 2016; Darling-Hammond, 2010). Accordingly, the income gap between individuals earning a college degree versus those who had earned a high school diploma nearly doubled from 1979 and 2012 (Autor, 2014). Arguably, the value of earning a college degree has never been higher due to the widening income gap.

Income inequity is merely a symptom of a greater problem. Access to high quality education programs at the K–12 levels provides the opportunity for students to lift themselves out of poverty (Darling-Hammond, 2010). However, Darling-Hammond (2010) argued that a student's race, ethnicity, and socioeconomic status largely determine educational opportunity (Darling-Hammond, 2010). Transitively, future income becomes a function of race, ethnicity, and socioeconomic status. Therefore, it seems logical to conclude that an increase in educational opportunity for at risk populations should erode the income inequity attributable to race, ethnicity, and socioeconomic status.

One measure of educational opportunity at the high school level is access to Advanced Placement programs (Cisneros, Gomez, Powers, Holloway-Libell, and Corley, 2014; Kettler & Hurst, 2017; Moller, Stearns, Southworth, & Potochnick, 2013). Researchers have found that

Advanced Placement programs increase the likelihood that students will enroll in college, graduate from college, and graduate from college in four years (Ackerman, Kanfer, & Calderwood, 2013; Bumpous, 2015; Chajewski, Mattern, & Shaw, 2011; Kettler & Hurst. 2017; Mattern, Marini, & Shaw, 2013; Shaw, Marini, & Mattern, 2012). Moreover, the inequity in educational opportunity extends to Advanced Placement programs in general and to Advanced Placement Calculus in particular (Cisneros et al., 2014; Handwerk, Tognatta, Coley, & Gitomer, 2008; Holmes, 2013; Kettler & Hurst, 2017; Kolluri, 2018; VanSciver, 2006). This lack of equity in access to Advanced Placement Calculus across the United States represents a threat to educational opportunity, which expands the income gap (Buckley, 2010; Darling-Hammond, 2010; Handwerk et al., 2008).

Background, Context, History, and Conceptual Framework for the Problem

The Advanced Placement program has seen steady growth since its inception, as has the Advanced Placement Calculus program (Holmes, 2013; Judson, 2017; Judson & Hobson, 2015). The Calculus AB exam is a widely utilized method for high school students to attain college credit for their high school mathematical learning (Judson, 2017; Judson & Hobson, 2015). However, researchers have found that there has is a distinct lack of access for Black and Hispanic students, as well as for students from homes that have low socioeconomic status, when it comes to the Calculus AB exam (Handwerk et al., 2008; James, Butterfield, Jones, & Mokuria, 2017; Kolluri, 2018; Moore & Slate, 2008; Walker & Pearsall, 2012). In addition to the problem of access to the Calculus AB exam, students who are Black or Hispanic face an ethnic achievement gap between themselves and their counterparts who are Asian or White (Holmes, 2013; Jara, 2013). A similar gender achievement gap has developed between males and females on the Calculus AB exam (Moore, Combs, & Slate, 2012; Morris, 2013; Morris & Slate, 2012).

Moreover, there has been a historical achievement gap between students of various ethnicities on Advanced Placement exams (Holmes, 2013; Jara, 2013; Moore & Slate, 2010). Researchers have even found that student performance on the Calculus AB exam is a function of students' state of residency (Davis, 2012; Koch, 2012; Wilson, 2013). Therefore, despite the nearly exponential growth in Advanced Placement participation, inequities in the form of achievement gaps persist.

Nationally, the problem of Advanced Placement equity includes three different facets. First, Advanced Placement programs are less likely to exist at schools that are predominantly Black, Hispanic, or low socioeconomic status (Handwerk et al., 2008; Kolluri, 2018; Moore & Slate, 2008; VanSciver, 2006). Second, even when students who are Black, Hispanic, or low socioeconomic status are afforded the opportunity to take Advanced Placement classes, those students usually score lower than their counterparts (Handwerk et al., 2008; Klopfenstein, 2004a; Moore & Slate, 2008; VanSciver, 2006). Third, male students consistently outperform female students on most Advanced Placement exams, including the Calculus AB exam (Melsom, 2008; Morris, 2013; Moore et al., 2012; Morris & Slate, 2012).

Interstate comparisons have also uncovered differences in Advanced Placement exam scores for students of the same ethnicity (Davis, 2012; Koch, 2012; Wilson, 2013). Although researchers have theorized about reasons for these differences, the reason that students of the same ethnicity consistently score higher in certain states is unclear. Koch (2012) theorized that preparedness and aligned curriculum in prerequisite classes might explain some of the difference. However, the literature seems to lack comparative analysis of state education policies as a possible contributor to Advanced Placement exam achievement gaps.

One example of how state policy may affect students' achievement on their Advanced Placement exams occurs in Michigan. In Michigan, state policy creates a temporal equity issue between students in the state of Michigan and their peers in other locales. The temporal inequity in Michigan is a result of the bookends of a legislated mandatory statewide start of school after Labor Day and College Board's requirement that nationwide Advanced Placement testing occur simultaneously each May (College Board, 2017b; Rowland, 2014). In comparison, many states allow individual school districts to set their own start dates, while other states simply direct that school cannot start before August 1, August 15, or the last week of August (Rowland, 2014). The results of these constraints is that students in Michigan start up to five weeks behind their counterparts in other states and frequently still have a month (or more) of school remaining after the Advanced Placement exam is taken, which cannot be used to prepare students for the aforementioned exams.

The issue of sufficient instructional time for Advanced Placement students may be a significant factor for first and second-generation immigrant students. Students whose parents do not speak English may be required to translate for parents at the family's medical, dental, or legal appointments (Bauer, 2016; Guan, Greenfield, & Orellana, 2014; Katz, 2014). Families of students who are first or second-generation immigrants often take family vacations to visit relatives still living in their native land during the school year and these vacations may extend well past the allotted vacation time by the school district (Sánchez, 2014; Urrieta, 2016). Some parents remove their children from school and allow them to attend political protests of immigration policy (Bolaños-Lopez, 2017, Cashmere, 2018; Gonzales, 2017). These familial commitments and parental requirements likely contribute to some of the achievement gaps that researchers have documented.

In order to address the impact of time on the learning of Advanced Placement Calculus students, the conceptual framework used in this study was opportunity to learn (OTL) theory. Researchers have frequently applied OTL theory to the linguistic acquisition of English Language Learners, which is particularly salient in a largely Hispanic school setting (Aguirre-Muñoz & Amabisca, 2010; Abedi & Herman, 2010). Due to the combination of temporal issues facing students in the state of Michigan, especially recent immigrants, this study focuses on the temporal variable.

Statement of the Problem

Mathematics achievement is a correlate of future earnings; therefore, it is essential to provided equitable opportunity to students of all genders, ethnicities, and socioeconomic backgrounds to ensure that they have the opportunity to pursue their career aspirations (Crawford & Cribb, 2013; James, 2013). However, researchers have found a gender achievement gap pertaining to the Calculus AB exam in various locales, which seems to affect students of every ethnicity (Holmes, 2013; Morris, 2013). These achievement gaps translate into an opportunity gap in collegiate choice, graduation rate, and collegiate achievement (Bumpous, 2015; Kettler & Hurst, 2017; Keng & Dodd, 2008; Mattern et al., 2013). However, there is an opportunity gap in K–12 education for Black, Hispanic, and low socioeconomic status students; and, this opportunity gap results in an achievement gap on the Calculus AB exam (Handwerk et al., 2008).

More concisely, the problem investigated within this study was the ethnic and gender achievement gaps on the Advanced Placement Calculus exam. While there has been considerable research demonstrating that these gaps exist, researchers have published very little research attempting to address the problem. This study considered whether increasing the length of the instructional period could reduce or eliminate the ethnic and gender achievement gaps on the Calculus AB exam.

Purpose of the Study

The purpose of the this study was to determine whether increasing the amount of instruction in Advanced Placement Calculus classes would have an effect on either the gender or ethnic achievement gaps on the Calculus AB exam. To accomplish this, it was necessary to determine whether ethnic and gender differences exist in Calculus AB exam achievement in the state of Michigan. As the current study included two components, the bifurcated investigation of ethnic and gender differences included both an achievement gap analysis and a quasi-experimental component to investigate whether an increase in the amount of instruction might be a strategy that educators can employ to reduce the size of the achievement gap. Potential benefits of the research are the uncovering of a scheduling technique that could provide equity to Black, Hispanic, and female students, while potentially saving those students both money and the trouble of taking a math placement exam upon enrollment in college.

In summary, this study was an attempt to address the ethnic achievement gap pertaining to the Calculus AB exam. The main purpose of the study was to determine whether increasing instructional time from one hour to two hours might solve the achievement gap problem that plagues the Calculus AB exam. After verification of the existence of gender and ethnic achievement gaps on the Calculus AB exam through an achievement gap analysis, the focus of the investigation shifts to a quasi-experimental component to determine whether a simple schedule change might offer educators a possible solution to the problem.

Research Questions

The following research questions were the central focus of this investigation as to whether gender and ethnic differences exist in Calculus AB exam achievement in Michigan and whether increasing the amount of instruction that students receive might be a potential remedy to reduce the size of any achievement gaps that still exist.

1. What is the difference in the Calculus AB exam performance of Michigan students as a function of gender?

2. What is the difference in the Calculus AB exam performance of Michigan students as a function of ethnicity?

3. What is the difference in the Calculus AB exam performance of Michigan's Hispanic students as a function of gender?

4. What is the difference in Calculus AB exam performance of students as a function of a yearlong two-hour AP Calculus class versus a one-hour AP Calculus class?

5. What is the difference in Calculus AB exam performance of Hispanic students as a function of a yearlong two-hour AP Calculus class versus a one-hour AP Calculus class?

6. What is the difference in Calculus AB exam performance of Hispanic male students as a function of a yearlong two-hour AP Calculus class versus a one-hour AP Calculus class?

7. What is the difference in Calculus AB exam performance of Hispanic female students as a function of a yearlong two-hour AP Calculus class versus a one-hour AP Calculus class?

Rationale, Relevance, and Significance of the Study

Researchers have performed many investigations and studies to determine whether there are gender or ethnic achievement gaps on the Calculus AB exam. Researchers have also conducted studies in an effort to determine whether student achievement on the Calculus AB exam is a function of state of residency. However, most of the research on achievement gaps on the Calculus AB exam is at least five years old at this point, and research using only students in the state of Michigan is seemingly nonexistent.

Researchers have also investigated the effects of various schedules on Calculus AB exam scores. However, most of the research on scheduling is two decades old. The intention of much of the research on school schedules versus Calculus AB exam performance was to inform schools and districts on how to organize school schedules to maximize Advanced Placement exam scores. These studies did not investigate achievement differences between boys and girls, nor were achievement gaps between different ethnicities examined.

While researchers have uncovered both gender and ethnic achievement gaps on the Calculus AB exam and some researchers have published theories about why these achievement gaps exist, there seems to be very little research discussing how schools and districts can solve the achievement gap problem. Consequently, the goal of this study was to test whether a change in the school schedule, specifically an increase of daily instructional time from one hour to two hours for AP Calculus, might offer a means for educators to remedy the twin gender and ethnic achievement gaps on the Calculus AB exam.

Definition of Terms

In order to instill clarity, this section lists definitions for recurring terminology. Presented in this section are operational definitions pertinent to the conceptual framework of the study, the rationale of the study, and the variables examined within the study.

Achievement gap. An achievement gap is a difference between the overall statewide average and the corresponding statewide average for a subgroup, or alternately, the difference between a given subgroup and the highest performing corresponding subgroup (U.S. Department of Education, 2012).

Advanced Placement (AP) courses. Advanced Placement courses are high school courses designed to follow a college-level curriculum designed by College Board (2016a). Each teacher of an Advanced Placement course must submit a course syllabus and have it approved by College Board (2016a) prior to labeling the course as Advanced Placement or AP.

Advanced Placement Calculus AB course. Advanced Placement Calculus AB course is a course designed by College Board (2016a). The course outline for Advanced Placement Calculus AB course covers the material typically taught during the first semester of a college Calculus course (College Board, 2016a).

Advanced Placement exam. In the context of this study, an Advanced Placement exam is any exam, designed by College Board (2016a), which students typically take after the successful completion of an Advanced Placement course to attempt to earn college credit.

Asian. In the context of this study, the term Asian refers to a student or group of students who have self-identified as Asian by marking that ethnicity on their Advanced Placement exam answer document.

Black. In the context of this study, the term Black refers to a student or group of students who have self-identified as Black by marking that ethnicity on their Advanced Placement exam answer document.

Calculus AB exam. In the context of this study, Advanced Placement Calculus AB exam is an exam designed by College Board (2016a) that students take to attempt to earn college credit for the Calculus course that they have successfully completed.

College Board. In the context of this study, College Board (2016a) is the non-profit entity that oversees the Advanced Placement program. Consequently, College Board (2016a) develops the Advanced Placement curriculum and designs the Advanced Placement exams.

Content exposure. In the context of this study, content exposure refers to the duration or amount of time that a teacher or school provides students to learn a particular concept or skill (Elliott & Bartlett, 2016).

Female. In the context of this study, the term female refers to a student or group of students who have self-identified as female by marking that gender on their Advanced Placement exam answer document.

Hispanic/Latino. In the context of this study, the term Hispanic refers to a student or group of students who have self-identified as Hispanic by marking that ethnicity on their Advanced Placement exam answer document.

Male. In the context of this study, the term male refers to a student or group of students who have self-identified as male by marking that gender on their Advanced Placement exam answer document.

White. In the context of this study, the term White refers to a student or group of students who have self-identified as White by marking that ethnicity on their Advanced Placement exam answer document.

Assumptions

In the course of this study, there were three assumptions. The first assumption was that any curricular reorganizations that the instructor made had a negligible impact on student achievement. The second assumption was that the quality of content delivery by the instructor produced a negligible effect on student achievement. The third assumption was that the number of days of actual instruction was approximately equivalent for each school year under consideration, as the number of instructional days subtracted from the school calendar for snow days, field trips, emergency drills, assemblies, and mandatory state and district testing days was unavailable for each school year.

Delimitations

The delimitations of this study included the following four different restrictions. First, the study was delimited geographically. Therefore, the delimitation on the achievement gap analysis was to the state of Michigan, while the delimitation on the quasi-experimental component was to a single Michigan high school. Second, the study was delimited temporally, as the achievement analysis portion of the study was delimited to the years 2002, 2007, 2012, and 2017, whereas the quasi-experimental component was delimited to the years 2014 through 2017. Third, among all Advanced Placement exams, the Calculus AB exam was the only exam included in this study. Fourth, the study included measurement of only a single dimension of OTL, which is instructional time.

Limitations

There were five limitations identified in the course of this study. The first limitation of this study is that the results of this study may not translate to other locales. The second limitation is that the conclusions obtained in the quasi-experimental component are constrained by the accuracy of the concordance tables used to determine the statistical equality of the comparison and experiential groups. The third limitation of this study is that measurements of student achievement are only as accurate as the instrumentation as each students only took the Calculus AB exam once. The fourth limitation was the inability of the metrics used in this study to measure cultural effects and their impact on student achievement. The fifth limitation was the inability to measure the effect of any dimensions of OTL on student achievement aside from the key concept of instructional time.

Summary

Researchers have determined that income gaps are the result of educational opportunity gaps. One of the many educational gaps is in equity of access and opportunity to Advanced Placement programs. The Calculus AB exam is an exemplar of the access and opportunity gaps that occur in the Advanced Placement programs, as there is a discrepancy in equity of access and achievement between students of various ethnicities. An achievement gap also exists on the Calculus AB exam between male students and their female counterparts.

In response to the various achievement gaps for Advanced Placement Calculus students, this study is an investigation into the question of whether OTL Theory might offer a remedy to the problem. Specifically, the primary purpose of this study was to determine whether increasing the length of the instructional period might reduce the gender and ethnic achievement gaps on the Calculus AB exam. This investigation is significant because research on how teachers should address the achievement gaps on the Calculus AB exam is virtually nonexistent.

Chapter 2: Literature Review

Two measures of college success are GPA and degree attainment rate. High school students who successfully complete an Advanced Placement course are more successful in both measures than their peers who do not take Advanced Placement courses, as the former group achieves higher college GPAs and are both more likely to attain a college degree in four and in six years (Ackerman et al., 2013; Bumpous, 2015; Chajewski et al., 2011; Kettler & Hurst, 2017; Mattern et al., 2013; Shaw et al., 2012). Thus, the Advanced Placement achievement serves as a leading indicator of college success (Bumpous, 2015; Keng & Dodd, 2008; Mattern et al., 2013; Shaw et al., 2012). Consequently, Advanced Placement achievement gaps pose a threat to the educational futures of high school students. Additionally, one key reason that the twin gender and ethnic achievement gaps in mathematics are not trivial is because mathematics skill is both a contributing factor to college readiness and a predictor of lifetime earnings (Altonji, Blom, & Meghir, 2012; Conley & French, 2013; Kim, Tamborini, & Sakamoto, 2015).

The Calculus AB exam annually exhibits both gender and ethnic achievement gaps (College Board, 2014; Holmes, 2013). The literature is replete with studies that show that boys outperform girls on math tests, including the Calculus AB exam, and with works that illustrate the achievement differences of various ethnic groups on the Calculus AB exam (Holmes, 2013; Inzlicht & Schmader, 2012; Jara, 2013; Moore & Slate, 2010, 2011; Morris, Slate, & Moore, 2015; Niederle & Vesterlund, 2010). While the primary focus of this study is narrowing, or ideally eliminating, the ethnic achievement gap on the Calculus AB exam, the two achievement gaps remain inextricably intertwined.

In the United States, the gender achievement gap in mathematics has a long, and oftentimes contentious, history. The stereotype is that males take more math classes than their female counterparts and males demonstrated higher achievement in mathematics than their female counterparts do (Else-Quest, Hyde, & Linn, 2010; Hyde, 2016). However, evidence supports the conclusion that the differences in mathematics between genders has narrowed in recent decades (Else-Quest et al., 2016). Niederle and Vesterlund (2010) found that males are only outscoring females by small amounts on achievement tests. Additionally, not every country exhibits a gender gap on standardized tests, such as Trends in International Mathematics and Science Study [TIMSS] and Program for International Student Assessment [PISA] (Else-Quest et al., 2010). This conclusion implies that the gender achievement gap may be due to cultural and environmental factors, rather than biological differences. In contrast, Fryer and Levitt (2010) assert that in the United States, a gender achievement gap in mathematics typically emerges during elementary school, it occurs at nearly every socioeconomic level, and it transcends ethnicity. A gender achievement gap in mathematics testing continues into high school and plagues Advanced Placement courses, as well as supplemental math contests (Ellison & Swanson, 2010; Moore et al., 2012; Morris & Slate, 2012).

Similarly, the ethnic gap in student achievement has a long complicated history in the United States. No Child Left Behind [NCLB] intended to stamp out the achievement gap that existed between White students and their Black and Hispanic peers (Reardon, Greenberg, Kalogrides, Shores, & Valentino, 2012). However, Black and Hispanic students continue to lag behind their Asian and White peers on achievement tests (College Board, 2014; Davis, 2012; Holmes, 2013; Wilson, 2013). The ethnic gap is not limited to student achievement alone, but extends to dropout and graduation rates, as well as college enrollment rates (Lewis, Simon, Uzzell, Horwitz, & Casserly, 2010). Consequently, Reardon et al. (2012) concluded that NCLB had mixed results in closing the ethnic achievement gap. In mathematics at the high school level, there is evidence that the ethnic achievement gap on the Calculus AB exam has not been closed (College Board, 2014).

Solutions to gender and ethnic achievement gaps seem to be less prevalent in the literature than studies whose subject is the existence of either achievement gap. College Board (2014) notes that only one state has closed the Advanced Placement gender gap. Additionally, both the achievement and participation gaps on the Calculus AB exam seem to be widening, instead of narrowing, with the rapid expansion in participation of said exam (Holmes, 2013; Jara, 2013; Moore & Slate, 2008; Walker & Pearsall, 2012). Solutions to the achievement gaps are rare and elusive within the literature. Moreover, teachers urgently need solutions to both gender and ethnic achievement gaps, especially in communities that are minority dominated, which are typically urban areas of low socioeconomic status. However, the efforts of Escalante, the legendary Bolivian Calculus teacher at Garfield High School in the seventies and eighties, may provide evidence of strategies that teachers, schools, and districts can use to mitigate the effects of urbanity and the poverty that frequently afflicts Hispanic communities.

Accordingly, the organization of this chapter includes six sections, beginning with this introduction. This chapter continues with a brief history of the Advanced Placement program, the conceptual framework of this study, and a review of research literature. This chapter concludes with a critique of previous research and a summary of Chapter 2.

Table 1

Recent Growth of Advanced Placement

Year	All Exams Taken	% of Increase (All Exams)	Calculus AB Exams Taken	% of Increase (Calculus AB)
2000	1,272,317	n/a	137,276	n/a
2005	2,105,803	65.5	185,992	35.5
2010	3,213,225	52.6	245,867	32.2
2015	4,704,980	46.4	308,215	25.4

Brief History of Advanced Placement

Advanced Placement tests have been given since the 1950s (Moore et al., 2012; Morris et al., 2015; Smith, Hurwitz, & Avery, 2017). The Advanced Placement program arose out of the Cold War to provide access to advanced levels of content for students while they were still in high school (Morris et al., 2015). College Board has administered the Advanced Placement program since 1955 (Morris et al., 2015). Since its inception, the Advanced Placement exam offerings have increased from 11 to 34 (Morris et al., 2015; Smith et al., 2017). Scoring of the Advanced Placement exams consists of mapping raw scores into integral scores ranging from a low of 1 to a maximum of 5 (College Board, 2016a; Keng & Dodd, 2008; Morris et al., 2015; Smith et al., 2017). Students may receive college credit for scores of 3 or higher, although each university establishes its own standard (Keng & Dodd, 2008; Moore et al., 2012). Early on, participation in the Advanced Placement testing was modest; however, during the 1990s, the number of students participating in Advanced Placement testing exploded (Klopfenstein, 2004b; College Board, 2015a). Consequently, the 490,299 Advanced Placement exams given in 1990 increased by 159% to 1,272,317 exams in 2000 (College Board, 2015a). As shown in Table 1, the subsequent decades witnessed continued growth in the Advanced Placement program. For example, the number of Advanced Placement exams College Board (2015b) administered in 2005 more than doubled by 2015. Similarly, the number of students sitting for the Calculus AB exam increased by a factor of approximately two-thirds between 2005 and 2015 (College Board, 2015b). Thus, the last two decades have witnessed exponential growth in Advanced Placement testing in general and in the Calculus AB exam in particular.

The increase in participation in Advanced Placement courses and exams reflects College Board's (2014) emphasis on equity. However, researchers have uncovered testing patterns that highlight equity concerns in the Calculus AB exam (Davis, Slate, Moore, & Barnes, 2014; Holmes, 2013; Moore & Slate, 2011; Morris, 2013; Morris et al., 2015). Some of these issues include discrepancies in scores between boys and girls and between the higher scoring Asian and White subpopulations and their Black and Hispanic counterparts (Holmes, 2013; Moore & Slate, 2011; Morris et al., 2015; Morris & Slate, 2012). Furthermore, 49 states and the District of Columbia continue to exhibit equity gaps for Black students, while 37 states still have an equity gap for Hispanic students (College Board, 2014).

Conceptual Framework

The conceptual framework for this study was OTL theory. The original definition of OTL stated that, "Opportunity to learn is defined as the amount of time allowed for learning, for example by a school schedule or program" (Carroll, 1989, p. 26). Since its inception, various researchers have revised the components of OTL theory, devised methods of measuring OTL, and further developed the theory itself (Elliott & Bartlett, 2016). The subsequent paragraphs in this section contain a brief history of OTL, current theory of measurement, and a summary of the critical ideas in this section.

A brief history of opportunity to learn theory. Carroll first introduced OTL in his presentation of the Carroll Model of School Learning in 1963, but he primarily developed the model to address the learning of language (Abedi & Herman, 2010; Carroll, 1989; Elliott & Bartlett, 2016; Wijaya, Van den Heuvel-Panhuizen, Doorman, & Veldhuis, 2018). Since that time, teachers and researchers continued to apply OTL to the learning of language; however, researchers have adapted the model for other content areas and applications, including special education and mathematics (Abedi, Courtney, Leon, Kao, & Azzam, 2006; Abedi & Herman, 2010; Walkowiak, Pinter, & Berry, 2017; Wijaya et al., 2018). The Carroll Model of School

Learning included five variables organized into two categories, namely instructional and individual (Elliott & Bartlett, 2016; Witt, Ulmer, Burris, Brashears, & Burley, 2014). Carroll believed that together the five variables, which are aptitude, OTL or time allowed for learning, perseverance, quality of instruction, and aptitude, or ability to understand instruction, accounted for the variance in student achievement (Elliott & Bartlett, 2016; Witt et al., 2014).

Since its introduction over half a century ago, researchers have continued to examine, develop, and refine OTL theory (Aguirre-Muñoz & Amabisca, 2010; Abedi & Herman, 2010; Elliott & Bartlett, 2016; Walkowiak et al., 2017). Some theorists believe that the variable of social support should be included in OTL, arguing, for example, that parental encouragement and expectations can provide additional OTL (Floden, 2002). In contrast to Carroll's Model that included five variables, Stevens and Grymes identified four variables that contributed to OTL (Cobb & Russell, 2015; Elliott & Bartlett, 2015; Witt et al., 2014). The four variables that Stevens and Grymes incorporated into their model were content coverage, content exposure, content emphasis; and quality of instructional delivery (Cobb & Russell, 2015). Through their definition of OTL, Stevens and Grymes highlighted the role of the teacher in the instructional process and underscored the importance of that role in maximizing student achievement (Cobb & Russell, 2015). Accordingly, OTL theory has seen consistent application with English Language Learners, as these students present a unique challenge to educators (Aguirre-Muñoz & Amabisca, 2010; Abedi & Herman, 2010). In their refinement of OTL mathematics, Walkowiak et al. (2017) alliteratively define OTL to include four dimensions. The Walkowiak et al. (2017) model includes the teacher's mathematical knowledge for teaching, time, tasks, and talk. Schools and districts have focused on the variables in OTL theory over which education providers have control; consequently, current theory frequently narrows the scope of OTL to three domains, which are instructional time, content, and quality of instruction (Elliott & Bartlett, 2016).

Measurement of opportunity to learn. Floden (2002) argues that the evidence in favor of OTL affecting student achievement is so convincing that researchers should shift their efforts towards improving the measurement of the construct itself. However, OTL is difficult to measure and since a universal metric does not exist, researchers utilize methods designed to focus on the variables they are emphasizing in their studies (Elliott & Bartlett, 2016; Wijaya, Van den Heuvel-Panhuizen, & Doorman, 2015). The Carroll Model distilled all of the variables used in the quotient into measures of time (Witt et al., 2014). Carroll expressed the degree of learning as a function of the quotient of time spent learning and the time needed by the student to learn (Witt et al., 2014). In the function, time spent learning is expressed as a product of the time allocated for learning and the perseverance of the learner while time needed to learn is expressed as a product of the aptitude of the learner, the quality of instruction, and the ability of the learner (Floden, 2002). However, certain components of OTL, such as time needed to learn and perseverance of the learner, remain difficult to quantify (Floden, 2002).

Boykin and Noguera (2011) believe that researchers should not measure instructional time by the amount of time that a teacher spends teaching, or even the amount of class time spent engaged with the curriculum. Instead, Boykin and Noguera (2011) conclude that instructional time equates to the amount of time that students spend performing various academic tasks, such as reading or performing math. Researchers have developed and used various systems of logging to tabulate the allocation of time and content taught (Elliott & Bartlett, 2016). However, frequently researchers have omitted the measurement of instructional quality from these measurement systems (Elliott & Bartlett, 2016). Limitations of the logging system include teachers' self-reporting data, extensive amounts of training that are required, and reliability and validity of the instruments themselves is still in question (Elliott & Bartlett, 2016).

Opportunity to learn theory summary. In summary, Carroll first outlined OTL theory over fifty years ago, and since then theoreticians have developed many variations of the original model of OTL (Elliott & Bartlett, 2016; Walkowiak et al., 2017; Wijaya et al., 2015; Witt et al., 2014). Wider application of OTL theory, combined with increased accountability, spurred refinement of OTL through further research (Aguirre-Muñoz & Amabisca, 2010; Abedi & Herman, 2010). Some researchers have developed models that eliminated factors, while others added new variables (Elliott and Bartlett, 2016). However, the two components that transcend the various models are the variables of instructional time and content (Elliott & Bartlett, 2016). Unfortunately, measuring OTL is difficult and there is no conventional metric to quantify the variables within OTL (Elliott & Bartlett, 2016; Floden, 2002; Wijaya et al., 2015).

Review of Research Literature

This section of this chapter contains a review of the literature pertinent to the problem of achievement on the Calculus AB exam. This literature review begins with an examination of studies that addressed gender and ethnic differences in Advanced Placement Calculus and then continues with an investigation of the OTL dimension of instructional time and its impact on learning. Subsequently, this review of the research literature contains an analysis of studies that addressed the various factors that contribute to student achievement and an investigation of the strategies Escalante utilized to increase Calculus AB exam scores at Garfield High School. This section concludes with a brief summary of the major findings within the literature.

Gender studies. Gender inequity in both mathematics and the Advanced Placement exams have a long and contentious history in the United States (Morris, 2013). Consequently, the literature contains gender studies of various types, several of which are relevant to the gender gap exhibited by the Calculus AB exam. This subsection of the literature review discusses the possibility of activation of stereotype threat on Advanced Placement exams, investigates gender comparisons for various individual ethnic groups on Advanced Placement exams, and concludes with gender comparisons for the Calculus AB exam.

Stereotype threat occurs when awareness of a cultural stereotype contributes to realization of the stereotype (Ben-Zeev et al., 2017; Inzlicht & Kang, 2010; Kapitanoff & Pandey, 2017). When stereotype threat is present, it negatively affects women's performance in mathematics (Kapitanoff & Pandey, 2017; Pennington & Heim, 2016). Black students have been found to perform worse on challenging high stakes tests when compared to their White peers (Kellow & Jones, 2008). Steele and Aronson (1995) conducted a series of four experiments to test whether stereotype threat affected the performance of individuals that belonged to a stereotype vulnerable group. Steele and Aronson (1995) concluded that asking about gender and ethnicity could produce the lower test scores for Black students on the SAT. In an attempt to determine whether asking students to identify their race and gender on their Calculus AB exam answer documents activated stereotype threat, Stricker (1998) conducted an experiment using a parallel design with 1,652 subjects. In the control group, students answered questions about their gender and ethnicity before the exam, while the experimental group answered questions about their gender and ethnicity after the exam. Despite conducting 21 two-way comparisons between the subjects in the control group, Stricker (1998) found no statistically significant differences between the two groups. Consequently, Stricker (1998) concluded that identifying gender and race on an answer document before taking a Calculus AB exam did not activate stereotype threat.

Some recent studies of gender differences in achievement on Advanced Placement exams have been restricted to a single ethnicity. One such study, conducted by Moore and Slate (2010), examined the Advanced Placement scores of American Indian students. American Indian female students had a higher incidence, 28.03% versus 23.88%, of scoring a 1 on an Advanced Placement exam than did their male counterparts (Moore & Slate, 2010). American Indians of both genders were nearly equally likely to score a 3 on an Advanced Placement Exam, but males were almost twice as likely to score a 5, 8.06% versus 4.18% (Moore & Slate, 2010). In their overall analysis of American Indian Advanced Placement scores, Moore and Slate (2010) found that there was a statistically significant difference between the distribution of American Indian males and their female counterparts. The effect size reported by Moore and Slate (2010) was .11, which the researchers interpreted as a small effect. Thus, American Indian female students were more likely to achieve the two lowest scores and less likely to score either of the two highest scores than their male counterparts.

Moore and Slate (2011) conducted a similar study in which they studied the Advanced Placement scores of Asian students. In their study, Moore and Slate (2011) analyzed Advanced Placement exam scores from fourteen consecutive years. The researchers observed that Asian boys outscored Asian girls on Advanced Placement exams in every single one of the fourteen years that were included in their study (Moore & Slate, 2011). The results that Moore and Slate (2011) reported were statistically significant in every one of the fourteen years, with small effect sizes of either .07 or .08 in each year from 2002 to 2010. In the most recent year of the study, that is 2010, Moore and Slate (2011) found that 24.91% of Asian boys achieved the highest score of 5 on their Advanced Placement exams, whereas 19.47% of Asian girls scored a 5. By a similar margin, Asian girls were more likely than their male counterparts (36.50% versus

31.27%) to score either a 1 or a 2 on their Advanced Placement exams (Moore & Slate, 2011). It is noteworthy that Moore and Slate (2010, 2011) examined scores from multiple Advanced Placement exams, and not just those pertaining to Calculus.

Other studies investigated gender achievement on Advanced Placement test for Black students. One study showed that in 2004, females accounted for 65.5% of Black Advanced Placement test takers compared to only 34.5% males (JBHE Foundation, 2004). Despite the discrepancy in test participation, on average Black males outscored their female counterparts, as the mean for males was 2.20 compared to 2.06 for females (JBHE Foundation, 2004). JBHE Foundation (2004) did not determine whether the differences between the genders were statistically significant, so they did not report effect sizes in their study. Finally, JBHE Foundation (2004) did not disaggregate the results reported in their study of Black Advanced Placement exam scores by specific test; however, they did report that the difference in mean scores was attributable to math and science exams.

In a nationwide statistical comparison of Asian, Black, Hispanic, and White ethnicities, Holmes (2013) chose to disaggregate her results by gender. Holmes (2013) utilized the mean score for each ethnicity and gender to build a stair step of the Advanced Placement results of the eight groups, using one group for each gender for each of the four ethnicities. Holmes (2013) found that among Asian, Black, and White students, male students outperformed their female counterparts in every single year from 1997 to 2012. In 2011 and 2012, Hispanic males outscored Hispanic females also, as evidenced in Table 2 (Holmes, 2013). In the ordering component of her study, Holmes (2013) did not indicate statistical significance or effect size. Additionally, Holmes (2013) did not disaggregate results by specific test.

Table 2

Mean Scores by Ethnicity and Gender for 2012

Ethnicity	Female	Male
Asian	3.14	3.33
Black	1.96	2.06
Hispanic	2.34	2.43
White	2.95	3.17

Other studies on gender differences have isolated Calculus AB exam data. One longitudinal analysis of Calculus AB exam scores, a study that used data from between 1996 and 2007, demonstrated that there was a statistically significant difference in participation in favor of males on the Calculus AB exam (Melsom, 2008). Additionally, Melsom (2008) concluded that males were outscoring females by statistically significant margin on the Calculus AB exam; although, in her study, Melsom combined data from all twelve years of the study to draw said conclusion. Another longitudinal study that analyzed data from the 2000, 2005, and 2010 Advanced Placement Calculus exams, found that boys scored statistically significantly higher on the Calculus AB exam in every year of the study, with effects sizes ranging from .10 to .13 (Morris and Slate, 2012). For the 2010 exam, Morris and Slate (2012) found that boys achieved the top mark of 5 at a higher rate than girls did, as 24.81% of boys compared to 16.18% of girls. Similarly, girls scored the lowest mark of 1 at a rate of 37.81%, while 29.61% of boys earned a 1 on the Calculus AB exam (Morris & Slate, 2012). In contrast to Moore and Slate (2010, 2011), Morris and Slate (2012) did not disaggregate gender differences by ethnicity.

Another longitudinal study examined Advanced Placement scores from 2007 and 2011 (Moore et al., 2012). In their study, Moore et al. (2012) concluded that boys were more likely than girls to take the Calculus AB exam, as in 2007, boys outpaced girls 105,008 to 99,538 in exams taken; and, in 2011, 126,484 boys took the Calculus AB exam compared to 119,088 girls.

In 2011, 16.81% of boys scored a mark of 5 on their Advanced Placement exam compared to 12.44% of girls (Moore et al., 2012). Additionally, 23.07% of girls who took an Advanced Placement exam scored the lowest mark of 1, whereas 19.33% of boys attained that same score (Moore et al., 2012). The combination of the pair of disparities, that is boys being more likely to achieve the highest mark and girls being more likely to score the lowest mark, resulted in a statistically significant difference in scores for boys and girls on Advanced Placement exams in 2011 (Moore et al., 2012). Moreover, boys scored better than their female counterparts did on the Calculus AB exam, although with a small effect size of .09 (Moore et al., 2012).

In another longitudinal study, Morris (2013) found that from 1997 to 2012, boys took more Calculus AB exams than girls did, with the narrowest margin in 2008, when boys held a 51.26% to 48.74% advantage, a discrepancy of 2.52%. Additionally, boys were more likely than girls to achieve a score of 5 on the Calculus AB exam in each of the sixteen consecutive years beginning in 1997 and ending in 2012 (Morris, 2013). Consequently, in 2012, the most recent year of Morris' (2013) study, 27.98% of boys earned a 5 on the Calculus AB exam compared to 21.12% of girls who earned the top mark. Notably, Morris (2013) reported small effect sizes for the difference in the distribution of Advanced Placement scores for boys and girls, with effect sizes ranging from a low of .07 in 2010 to a high of .12 in both 1998 and 2001.

In a meta-analysis of studies on gender difference in math and science, Morris et al. (2015) found convergent results, as boys outscored girls in every one of the eight studies they reviewed. Additionally, Morris et al. (2015) found similar convergence in their analysis of studies of gender differences on Advanced Placement exam scores in science and math, as boys outperformed girls by a statistically significant margin in the seven studies they examined. The two primary theories that explain the differences in achievement in mathematics are frequently labeled nature and nurture (Good, Aronson, & Harder, 2008; Morris et al., 2015). The first theory, the nature theory, posits innate differences between the two genders which results in higher achievement in mathematics for boys (Good et al., 2008; Morris et al., 2015). The second theory, the nurture theory, proposes that higher achievement for boys in mathematics is a result of the confluence of cultural, social, and environmental factors (Good et al., 2008; Morris et al., 2015). Researchers have incorporated some of the factors identified by the second theory into their OTL models, such as increased opportunities for students to talk (Walkowiak et al, 2017).

In summary, the studies examining gender differences yield comparable findings (Morris et al., 2015). Achievement differences exist between the two genders in Advanced Placement scores in the whole population, as well as in subpopulations including American Indian, Black, and Asian students (Holmes, 2013; JBHE Foundation, 2004; Melsom, 2008; Moore & Slate, 2010, 2011; Moore et al., 2012; Morris, 2013; Morris & Slate, 2012; Morris et al., 2015). While the statistical significance was not reported, boys are taking more Calculus AB exams than girls are every year (Melsom, 2008; Moore et al., 2012; Morris, 2013). Boys are more likely than girls to achieve a score of 5 on any Advanced Placement exam, and less likely to earn the lowest score of 1 (Moore et al., 2012; Morris & Slate, 2012). Furthermore, boys are more likely than girls to earn a score of 5 on the Calculus AB exam (Moore et al., 2012; Morris, 2013; Morris & Slate, 2012). In most studies, the study author characterizes the effect size for the Advanced Placement gender achievement gap as small. To synthesize these conclusions, more boys are taking the Calculus AB exam and, on average, boys are outperforming girls.

Ethnic studies. College Board (2014) has communicated a desire for equity in its Advanced Placement program. Despite the push for inclusiveness in the Advanced Placement program, only one state has managed to eradicate the equity gap for Black students and only thirteen have managed to accomplish the same feat for Hispanic students (College Board, 2014). Therefore, this section will examine the literature pertaining to equity in the Advanced Placement program for various ethnic groups.

Hispanic students face challenges to gaining access in Advanced Placement programs. In Texas, both Black and Hispanic students are underrepresented in Advanced Placement classes (Klopfenstein, 2004a). Although evidence of a disparity in the access to Advanced Placement courses in Texas was not present, both Black and Hispanic students opted not to take those courses (Klopfenstein, 2004a). A major difference in the population studied by Klopfenstein (2004a) involved students from low-income homes; a characteristic that handicapped Black and Hispanic students at rates of 3:1 compared to their White peers, but heavily influenced the likelihood of student participation in the Advanced Placement program. Klopfenstein (2004a) did not report whether the differences that she found were statistically significant.

In a study of Advanced Placement participation in the state of California, Solórzano and Ornelas (2004) identified the schools with the most Advanced Placement students per capita, and subsequently dissected the data emerging from these programs. Although Hispanic students constituted 38% of the high school population in the state of California, this ethnic group accounted for a mere 16% of students in the 50 schools with the most Advanced Placement courses per capita (Solórzano & Ornelas, 2004). Likewise, Hispanic students remain underrepresented in Advanced Placement courses in Arizona (Cisneros et al., 2014). Although Los Angeles Unified School District was 66% Hispanic during the 2001–2002 school year, only 49% of the students enrolled in Advanced Placement courses in the district were Hispanic (Solórzano & Ornelas, 2004). Consequently, Solórzano and Ornelas (2004) concluded that, in

California, there was a disproportionate underrepresentation of Hispanic students in the Advanced Placement program, even when Hispanics were a majority of the local population.

In a recent qualitative study, Walker and Pearsall (2012) studied a Western suburban high school whose Hispanic population was 62% of the student body, while just 14% of its Advanced Placement students were Hispanic. Walker and Pearsall (2012) concluded that there is a persistent underrepresentation of Hispanic students in Advanced Placement courses; and, Hispanic students underachieve in those courses. Through their qualitative study using focus groups, Walker and Pearsall (2012) determined that many of the reasons that Hispanic students chose not to take Advanced Placement courses were social and cultural. For instance, some Hispanic students reported that two such factors were avoidance of racial labeling and fear of failure (Walker & Pearsall, 2012).

In another qualitative study, VanSciver (2006) conducted interviews of low-income minority students and parents. One phenomenon that contributes to the Hispanic underrepresentation in Advanced Placement Calculus is that Hispanic students purposefully under schedule themselves in order to take home better grades (VanSciver, 2006). Another conclusion drawn by VanSciver (2006) was that parents implicitly trusted the school district and its personnel to manage the education of their children. The combination of these phenomena resulted in very few students entering the pipeline to Advanced Placement Calculus (VanSciver, 2006). Thus, to combat the lack of Hispanic participation in Advanced Placement, a Delaware district analyzed middle school test data to find students with academic talent, schedule them for rigorous courses, and support them with counseling and tutoring services (VanSciver, 2006). The program succeeded in increasing minority participation in the courses throughout the pipeline from Algebra 1 to Calculus from 2001, where only two minority students were in the pipeline, to

2004–2005, where 180 students were in the pipeline leading to Advanced Placement Calculus (VanSciver, 2006).

In similar longitudinal studies, Holmes (2013) and Jara (2013) contrasted both the Advanced Placement exam tendencies and scores of students by ethnicity. Jara (2013) found that, throughout a study that analyzed 13 years of data, Hispanic students were twice as likely to take the Advanced Placement English Language and Composition exam as the Calculus AB exam. In contrast, at the beginning of the study, White students were more likely to take the Calculus AB exam than the Advanced Placement English Language and Composition exam (Jara, 2013). By 2012, White students were still more than 1.5 times as likely to take the Advanced Placement English Language and Composition exam as the Calculus AB exam (Jara, 2013).

The differences between ethnicities were not just limited to participation rates. Asian and White students outperformed their Black and Hispanic counterparts every year between 1997 and 2012 (Holmes, 2013; Jara, 2013). For example, the mean scores on Advanced Placement exams in 2000 were 3.06 for both Asian and White student subgroups, while Black and Hispanic students posted mean scores of 2.77 and 2.17 respectively (Holmes, 2013; Jara, 2013). In 2012, the most recent year examined by Holmes (2013) and Jara (2013), the mean scores were 3.23 and 3.05 for Asian and White students respectively, meanwhile Black and Hispanic students scored means of 2.38 and 2.10 respectively (Holmes, 2013; Jara, 2013). The difference in the means was statistically significant in every comparison, indicating that Advanced Placement scores vary by race; although, effect sizes for those comparisons ranged from trivial to small (Holmes, 2013; Jara, 2013). Additionally, Holmes (2013) found that the ethnic scoring differences were statistically significantly different for both genders, with monotonically increasing effect sizes for both genders. For females, effect sizes for ethnic achievement differences in Advanced Placement scores increased from .10 in 1997 to .14 in 2012; while, for males, effect sizes for ethnic achievement differences increased from .08 in 1997 to .15 in 2012 (Holmes, 2013). Although the ethnic achievement differences for both genders only moved from the trivial to the small range, Holmes (2013) concluded that there is a growing threat of ethnic disparity in Advanced Placement scores.

However, discrepancies in Advanced Placement exam scores between ethnicities were not limited just to means. In 2012, 16.07% of White students scored a 5 on an Advanced Placement exam, while only 8.39% of Hispanic students scored that high; meanwhile, 35.26% of White students scored below a 3, while 58.98% of Hispanic students scored one of the two lowest marks (Holmes, 2013; Jara, 2013). Effect sizes for the differences between scores of the two ethnic groups ranged from .07 in 2000 to .17 in 2011, although after 2002, the lowest effect size increased to .10, which translate to a small effect size for every year from 2002 to 2012 (Jara, 2013). Hence, an achievement gap persists between White students and Hispanic students on Advanced Placement exams (Holmes, 2013; Jara, 2013).

Even though College Board (2014) documents the elimination of an equity gap within a few states for various ethnic groups, an interstate equity gap also plagues the Advanced Placement program. Multiple studies have shown that the state of residency is a factor in Advanced Placement exam scores (Davis, 2012; Koch, 2012; Wilson, 2013). Davis (2012) compared the Advanced Placement scores of Black students in Texas, New York, and Florida. Similarly, Wilson (2013) compared the Advanced Placement scores of Black male students in Connecticut, Florida, Maryland, Massachusetts, and Texas. In both studies, the researcher found that the state of residency produced a statistically significant difference in Advanced Placement exam scores for Black students (Davis, 2012; Wilson, 2013). Black students in New York scored higher than their Black peers in Texas and Florida by a statistically significant margin every year from 1997 to 2012, with trivial to small effect sizes ranging from .06 to .17 (Davis, 2012). Koch (2012) compared the Calculus AB exam scores for Hispanic students from California, Texas, and Arizona extending from 1997 to 2012. The state of residency also affected Advanced Placement exam scores for Hispanic students, as the difference in the scores for Hispanic students in the three states was statistically significant for each year of the study, with trivial to small effect sizes ranging from .07 to .14 (Koch, 2012). Therefore, the results reported by Koch, (2012) dovetail with those of Davis (2012) and Wilson (2013), suggesting that the state of residency seems to influence Advanced Placement exam scores for minority students.

In summary, the literature demonstrates an underrepresentation of both Black and Hispanic students in taking Advanced Placement exams (Klopfenstein, 2004a; Solórzano & Ornelas, 2004; Walker & Pearsall, 2012). Both Black and Hispanic students scored lower than their Asian and White peers do, and are both more likely to score a 1 and less likely to score a 5 (Holmes, 2013; Jara, 2013). Moreover, the state of residency is a factor in determining Advanced Placements scores for minority students (Davis et al., 2014; Koch, Slate, & Moore, 2014). Both ethnicity and state of residency have yielded small effect sizes in Advanced Placement exam scores (Davis, 2012; Jara, 2013; Koch, 2012). However, VanSciver (2006) found that an intentional plan could bolster minority participation in the Advanced Placement program. Thus, researchers have found ethnic achievement differences in Advanced Placement exam scores include differences in measures of central tendency and of extreme scores.

Instructional time. The literature explains that various factors affect students' Advanced Placement exam scores (Burney, 2010). Fixed factors, such as socioeconomic status and school size, account for 63.1% of the variance in Advanced Placement exam scores (Burney, 2010). However, usage of time is not a fixed factor. Escalante encapsulated this thought in his maxim, "Use time correctly, and it will be your friend; use it poorly, it will be the enemy" (Gilroy, 2006, p. 35). Moreover, instructional time and its usage for effective educational purposes is central to OTL, as the variable of content exposure refers to the amount of instructional time allocated to learning (Abedi et al., 2006; Elliott & Bartlett, 2016). Consequently, this section will examine the literature pertaining to allocation of time, with a particular emphasis on school schedules and lengths of periods of instruction for Advanced Placement Calculus.

The 1990s witnessed a transition for many school districts away from traditional schedules, which featured periods that ranged from 45 to 60 minutes in length, to block schedules, which utilized longer instructional periods that were typically 75 to 100 minutes in length (Walsh, 2011; Zelkowski, 2010; Zepeda & Mayers, 2006). Educational leaders theorized that increasing the length of classes could result in comparable, if not improved, student achievement (Keen, 1996; Walsh, 2011). Proponents of block schedules argued that there were fewer discipline issues, while opponents to block scheduling cited decreased student achievement on standardized assessments, including Advanced Placement exams (Zepeda & Mayers, 2006). The debates between the two factions precipitated many studies that investigated the effects of the various permutations of traditional and block schedules on student achievement.

Many of the block schedules implemented in the 1990s were 4x4 block schedules, where students took four classes each semester (Zelkowski, 2010; Zepeda & Mayers, 2006). On a 4x4-block schedule, students change classes at the end of the first semester, which means that students typically take math during only one of the two semesters (Zepeda & Mayers, 2006).

Despite all of the attention attributed to school schedules, Wallace (2013) found that a school's schedule did not significantly influence student achievement in high school Algebra 1 classes. Similarly, the structure of the school schedule did not significantly affect student achievement on Florida's end of course Algebra 1 exam, nor did the schedule affect student achievement for males, females, Black, or Hispanic students (Underwood, 2014). In both Virginia and Florida, students took the end of course assessment upon completion of the course of study (Underwood, 2014; Wallace, 2013). For states that utilize end of course examinations, block scheduling seemingly offers schools a viable alternative to traditional scheduling. Contrastingly, through his analysis of NAEP test scores, Zelkowski (2010) found that students on 4x4 block schedules lagged approximately two thirds of a year behind their counterparts on year-long traditional schedules.

The Advanced Placement program offers a unique challenge to schools using non-traditional schedules as all students nationwide take Advanced Placement exams in May, regardless of when their course of study concluded (College Board, 2017b; Gullatt, 2006; Hansen, Gutman, & Smith, 2000). At a high school in Maryland, usage of a 4x4 schedule produced a decline in Advanced Placement Calculus exam scores, while simultaneously resulting in 7% fewer scores of 3 or higher (Guskey and Kifer, 1995). In a meta-analysis of studies of school schedules, Gullatt (2006) concluded that semester block schedules could have adverse effects on Advanced Placement scores. On the Calculus AB exam, Smith and Camara (1998) found that students in traditional yearlong courses outscored students in semester-long block classes, as did students in yearlong extended periods. Likewise, Keen (1996) found that students who took a single semester block schedule Calculus class scored lower than their peers did whose Calculus class was yearlong. In their meta-analysis of the literature on block scheduling, Zepeda and Mayers (2006) reported mixed results as some authors reported traditionally scheduled students outperformed their block scheduled peers on the Advanced Placement exams, while other studies reported the exact opposite conclusion.

Many schools that featured block schedules limited Advanced Placement instruction to a single semester (Smith & Camara, 1999). By requiring a fall prerequisite course, a spring Advanced Placement course on a block schedule yielded higher Advanced Placement Calculus scores, while producing 21% more students who took the exam and an increase of 71% in students scoring 3 or higher (Hansen et al., 2000). Likewise, 81% of Calculus AB students in one district using a semester block schedule scored 3 or higher when those students took Calculus AB in the fall and subsequently took Calculus BC in the spring (Lightner, 2009). Therefore, continuous instruction in mathematics is advisable to achieve optimal results on Advanced Placement Calculus exams (Smith & Camara, 1998; Zelkowski, 2010).

In a meta-analysis of literature on learning time, Kidron and Lindsay (2014) concluded that increased learning time has the potential to be effective in urban settings. Additionally, that increased learning time is especially advantageous for students who were, "…performing below standards" (Kidron & Lindsay, 2014, p. 10). While effect sizes were small, increased learning time also produced higher mathematics achievement (Kidron & Lindsay, 2014). For optimal results, increased learning time programs should feature a certified teacher and students who are not a risk (Kidron & Lindsay, 2014).

In summary, the literature on block scheduling identifies Advanced Placement as a consistently problematic area (Gullatt, 2006; Guskey & Kifer, 1995; Keen, 1996; Smith & Camara, 1998). Semester block classes typically produce lower Advanced Placement test scores than do yearlong courses (Gullatt, 2006; Guskey and Kifer, 1995; Keen, 1996). However, some schools have achieved success through programs whereby students take Calculus, or a prerequisite math course, each semester (Hansen et al., 2000; Lightner, 2009). Additionally, schools that feature yearlong block schedule Calculus classes produce increased student achievement (Gullatt, 2006; Smith & Camara, 1998). Therefore, yearlong extended classes seemingly offer hope to schools that seek to mitigate the effects of urban poverty.

Student achievement. This section presents the literature that contributes toward student achievement from four different perspectives. The first perspective discussed below is factors affecting student achievement in general. The second perspective is correlates of mathematics achievement. The third perspective is factors that determine Calculus achievement. The fourth perspective is an investigation of factors that contribute to Advanced Placement achievement. Therefore, the next four paragraphs each contain discussion of one of the aforementioned four perspectives

The first perspective from which researchers have studied student achievement is general student achievement. Hattie (2008) and Marzano, Pickering, and Pollock (2001) conducted separate meta-analyses of the literature in search of factors that contribute toward student achievement. Hattie (2008) analyzed over 800 works and classified the factors affecting student achievement into six categories: "1. the child; 2. the home; 3. the school; 4. the curricula; 5. the teacher; 6. the approaches to teaching" (p. 31). From those six domains, Hattie (2008) found that the top five influences on student achievement were the student self-reporting grades (effect size of 1.44), Piagetian programs that emphasize thinking processes instead of outcomes (effect size of 1.28), providing formative evaluation (effect size of .90), micro teaching (effect size of .88), and acceleration (effect size of .88). Through an extensive meta-analysis, Marzano et al. (2001) analyzed pedagogical techniques that result in higher student achievement. Marzano et al.

(2001) identified eight high yield strategies that correlated with student achievement. The four strategies with the highest effect size were identifying similarities and differences, which had an effect size of 1.61; summarizing and note taking, which had an effect size of 1.00; reinforcing effort and providing recognition, which had an effect size of .80; and homework and practice, which had an effect size of .77 (Marzano et al., 2001).

Second, as Calculus is the capstone of high school mathematics, it is necessary to investigate factors in mathematics achievement. A very strong correlate of mathematics achievement is socioeconomic status (Hernandez, 2014). Tosto, Asbury, Mazzocco, Petrill, and Kovas (2016) concluded that affective characteristics, such as self-perceived ability and subject interest, were more important than classroom environment in mathematics achievement. Siegler et al. (2012) found that understanding of both fractions and division have long-term implications for mathematics success, not just Algebra, but advanced mathematics as well. Kitsantas, Cheema, and Ware (2011) were surprised to find that increasing percentages of homework time spent on mathematics correlated negatively with mathematics achievement. Lamb and Fullarton (2001) found that classroom composition practices greatly affected student achievement. On the Third International Mathematics and Science Study, students who were in high ability groups scored much higher than those who were not (Lamb & Fullarton, 2001). Clotfelter, Ladd, and Vigdor (2010) found that experience and certification of the classroom teacher correlate positively with achievement in mathematics.

Third, although Calculus is an advanced mathematics course, researchers have investigated specific factors that determine achievement at the Calculus level. Champion and Mesa (2016) have determined that, “U.S. students’ race, socio-economic status, prior mathematics knowledge, mathematics course placement, and mathematics self-efficacy each has an important role in the likelihood of completing calculus in high school" (p. 22). In contrast, Ahmad et al. (2017) determined that it was solely student effort that translated into student achievement in college level Calculus. Sadler and Sonnert (2016) found that college professors preferred that high schools emphasize understanding mathematics, whereas high school teachers emphasized sound pedagogy. Sahin, Cavlazoglu, and Zeytuncu (2015) found that flipping a college level Calculus classroom increased student quiz scores and students felt that the flipped classroom enhanced their learning. In a similar study, Albalawi (2018) found that flipped classrooms not only result in higher student achievement, but they can also provide support for at-risk students. Easey and Gleeson (2016) found that school officials, including teachers and administrators, served as gate keepers by directing students to take less challenging courses, failing to promote the benefits of studying Calculus, and discouraging capable students from taking Calculus when the student wasn't sure about future career aspirations.

Fourth, in search of factors that determine success on Advanced Placement exams, researchers have determined that such factors include affective characteristics of the student, measures of previous achievement, and school-controlled factors (Burney, 2010; Henry, 2008; Peterson, 1989). Peterson (1989), in search of predictors of Advanced Placement Calculus success, found that strong correlates of the Calculus AB scores included previous standardized test scores in Mathematics and earned grades in Mathematics classes taken before sitting for the Calculus AB exam. In a study, investigating the differences between students who excel in Advanced Placement and Honors courses and those who do not, Henry (2008) found that affective characteristics, like self-efficacy and motivation, differed between those who excelled and those who did not. Many of the sources of variance in student achievement on Advanced Placement exams are due to fixed factors (Burney, 2010). Those fixed factors included school size, socioeconomic status, and ethnicity (Burney, 2010). Using hierarchical linear regressions to identify sources of variance, Burney (2010) found that such fixed factors accounted for 80% of the variance in student achievement on Advanced Placement exams. However, Burney (2010) also found that there were sources of variance that were controllable, such as social opportunities, grouping of students, and curriculum (Burney, 2010).

In summary, factors affecting student achievement fall into six separate domains (Hattie, 2008). While Hattie's (2008) research identified students self-reporting their own grades as the highest correlate of student achievement, Marzano et al. (2001) found that the pedagogical technique of identifying similarities and differences had the highest effect size. Multiple researchers have concluded that Calculus achievement may be aided using a flipped classroom at the collegiate level (Albalawi, 2018; Sahin et al., 2015). Researchers have also investigated grades and scores that are predictive of Advanced Placement success, personal characteristics of successful Advanced Placement students, and school and district variables that contribute to Advanced Placement success (Burney, 2010; Henry, 2008; Peterson, 1989). Peterson (1989) found that Calculus AB exam scores correlate positively with both grades in math classes and standardized Mathematics test scores. Affective characteristics are different for students who achieve success on the Advanced Placement exams and students who do not (Henry, 2008). Burney's (2010) analysis noted that while 80% of the variance in student achievement on Advanced Placement exam is due to fixed factors, schools do have the ability to control some variables, such as the curriculum.

Escalante's strategies. During the 1980s, Garfield High School achieved enormous success on the Calculus AB Exam (Claudet, 2016; Mathews, 1988). Despite mixed success in prior years, Escalante's 1981-1982 Calculus class excelled on the Calculus AB exam (Claudet, 2016; Mathews, 1988). At the time, the population of Garfield High School was 95% Latino and 80% free or reduced lunch eligible, a conventional proxy for socioeconomic status (Mathews, 1988). By 1987, Garfield High School ranked fourth in the nation amongst public schools in the number of students taking the Calculus AB exam (Mathews, 1988). Therefore, the strategies that Escalante implemented merit analysis.

The backbone of Escalante's pedagogy was motivation and mastery learning (Julie, Mbekwa, & Simon, 2016; Knudson-Martin, 2011). In addition to high expectations and a safe learning environment, Escalante credited a supportive principal for the success of the mathematics program at Garfield High School (Gradillas & Jesness, 2010). However, providing students sufficient time to learn may have been the most important factor to the success that Escalante and his students realized in Advanced Placement Calculus (Gradillas & Jesness, 2010). Escalante and Dirmann (1990) cite the importance of scheduling students for continuous intensive mathematics training as a critical contributor to the success of the mathematics program at Garfield High School. Sometimes this meant that students took more than one math class during a semester (Escalante & Dirmann, 1990). Additionally, Escalante asked that students arrive at school before classes began, stay after school, or participate in Saturday study sessions (Claudet, 2016). Escalante demanded hard work from his students; thus, his students reported spending almost nine hours per week outside of class on Calculus (Kester, 1993). In addition, Escalante's students took two hours of math each day (Glennon & Mohler, 1999).

Through an emphasis on time spent on learning, Garfield High School gave the fourth most Calculus AB exams in the nation in 1987 (Mathews, 1988). Moreover, 66% of Escalante's students passed the Calculus exam that year (Mathews, 1988). While requiring students to arrive early, stay after school, or attend Saturday study sessions is beyond the scope of most public schools, asking students to take two hours of Calculus instruction daily is not.

Summary of the literature

Studies repeatedly demonstrate that Advanced Placement scores of boys are typically higher than scores of girls, especially in science and math (Holmes, 2013; Jara, 2013; Moore et al., 2012; Moore & Slate, 2010, 2011; Morris, 2013; Morris & Slate, 2012; Morris et al., 2015). Similarly, both Asian and White students score higher on Advanced Placement Calculus exams than either Black or Hispanic students (Holmes, 2013; Jara, 2013). Other studies show that the structure of a school's schedule can also influence student Advanced Placement scores (Gullatt, 2006; Guskey & Kifer, 1995; Hansen et al., 2000; Smith & Camara, 1998). Previous grades in mathematics and Advanced Placement scores may predict student success on the Calculus AB exam; although, student success seems to be affected by both individual affective characteristics, such as motivation, and school-controlled variables, such as curriculum (Burney, 2010; Henry, 2008; Peterson, 1989). Escalante was able to harness a combination of extended instructional time and motivation, which translated into stellar achievement for his Advanced Placement Calculus students (Escalante & Dirmann, 1990; Gradillas & Jesness, 2010; Julie et al., 2016; Kester; 1993; Mathews, 1988).

Critique of Previous Research

Many of the studies in the review of literature included an analysis of Advanced Placement scores. As College Board (2016a) explained, Advanced Placement scores are categories of scores. Although the categories are ordinal in nature, the majority of researchers have followed College Board's lead in reporting and analyzing means. However, when dissecting ordinal data, pure statisticians usually eschew usage of the mean as the primary measure of central tendency in favor of the median (Adams & Lawrence, 2015; Gravetter & Wallnau, 2009).

In the analysis of gender differences, Melsom (2008) tested her hypotheses with two-way ANOVA and two-sample t-tests. Adams and Lawrence (2015) posit that both the ANOVA two-way test and the two-sample t-tests require interval data. Thus, Melsom (2008) chose to compare the percentages of test takers who scored a 3 or higher for the two genders. This meant that for each gender, Melsom (2008) essentially grouped the participants into a group that passed and a group that failed. However, some schools do not award university credit for a Calculus AB exam score of a 3 or a 4 (University of Michigan, 2017). Similarly, since a 1 is the lowest possible score, a considerable disparity in knowledge may exist between students who score a 1 and those who score a 2. Additionally, Melsom (2008) grouped all twelve years of data together. Melsom's (2008) approach of combining data for twelve years obscures the effects from each year and trends in the size of the achievement gap between the two genders are lost.

Similarly, Morris (2013) conducted a comparison of Advanced Placement performance between the two genders. Instead of using ANOVA and t-tests, Morris (2013) compared the results of each year individually using a Chi-square test. Morris (2013) chose to group failures together into a single category, but kept potentially passing scores in separate categories. Holmes (2013) used the same grouping of scores to analyze the difference between scores of four different ethnicities. Likewise, in his three state comparison of Black achievement on Advanced Placement exams, Davis (2012) chose to combine scores of 1 and 2 into a single category, while keeping scores of 3, 4, and 5 separate. Thus, there is the potential for a significant disparity in knowledge between scores of 1 and 2, as students may get 0% of the Advanced Placement questions correct and still receive a score of 1. In contrast, a student who scores a 2 may be getting almost 50% of the questions correct, so the merging of the two lowest scores impedes the possibility of identifying differences in learning between the two scores.

Contrastingly, Wilson (2013) combined scores of 1, 2, and 3 together to form the category of no potential for college credit and combined the scores of 4 and 5 together to form the category for college credit. Wilson's (2013) categories minimize the value of a score of 3, for which various schools still award credit (Central Michigan University, 2015; Ferris State University, 2016; University of Michigan-Dearborn, 2017; Wayne State University, 2015). Wilson's (2013) interpretation potentially overvalues a score of a 4, which will not earn a student credit at many universities, including the University of Michigan (2017).

In their investigations into whether the state of residency affects Advanced Placement exam scores, Wilson (2013), Davis (2012), and Koch (2012) each selected a single ethnicity to investigate. Davis (2012) and Wilson (2013) examined whether the state of residency affects the Advanced Placement exam scores for Black males, while Koch (2012) studied the same phenomenon for Hispanic participants. Although the results of the three studies unanimously indicate that the state of residency affects the Advanced Placement exam scores for Black and Hispanic students, it remains unclear whether this result holds for Asian and White students, as none of the three researchers investigated the phenomenon with these subpopulations. Moreover, it is not clear whether the effects of state of residency are stronger for one ethnicity than it is for the others.

Seemingly, the majority of Advanced Placement studies rely on data from the College Board and feature quantitative analysis. Consequently, there are few empirical studies examining the effects of experimental treatments on Advanced Placement student achievement. In contrast, studies involving schedule structures are ubiquitous, but nearly all of them compare traditional schedules with semester-long block classes. The studies that do examine yearlong block courses indicate that yearlong schedules can produce increased student achievement on Advanced Placement exams (Smith & Camara, 1998; Zelkowski, 2010). Therefore, it seems logical to consider implementation of yearlong block instruction in an attempt to close the Advanced Placement achievement gaps, although this idea is seemingly absent in the literature.

Chapter 2 Summary

Chapter 2 contained evidence that the Calculus AB exam continues to exhibit both gender and ethnic achievement gaps (Holmes, 2013; Moore and Slate, 2010, 2011; Moore et al., 2012; Morris, 2013; Morris & Slate, 2012; Morris et al., 2015). Although stereotype threat could be a factor in gender achievement differences on the Advanced Placement exam, it is likely that the gender gap is the result of more than just activating stereotype threat through the seemingly innocuous student information section of the answer document (Stricker, 1998; Stricker & Ward, 2004). Consequently, the principal topic of this study was gender differences related to Advanced Placement exam scores, instead of stereotype threat activation through identifying personal data on answer documents. Since equitable educational opportunity requires eliminating both the gender and achievements gaps, the focus of this study was on a possible way to eliminated achievement gaps on the Calculus AB exam.

The conceptual framework of this study was OTL theory. Despite the emergence of various models of OTL over the course of the more than 50-year history of OTL, every model seems to include both instructional time and content (Elliott & Bartlett, 2016). Consequently, OTL theory suggests that increased time on task should foster improved student achievement (Boykin & Noguera, 2011; Carroll, 1989; Ottmar, Decker, Cameron, Curby, & Rimm-Kaufman, 2013). Researchers report that yearlong Calculus instruction yielded higher scores on Advanced

Placement exams than semester block schedules do (Hansen et al., 2000; Keen, 1996; Lightner, 2009). Furthermore, the literature indicates that continuous instruction for extended periods may increase Calculus AB exam scores (Smith & Camara, 1998; Zelkowski, 2010).

Hattie's (2008) six domains that affect student achievement dovetail with the Carroll Model of School Learning, as both researchers recognize that there are individual and instructional components that affect learning. The top five strategies that Hattie (2008) found in his meta-analysis were all pedagogical in nature. In contrast, Marzano et al. (2001) identified additional high-yield pedagogical strategies that teachers use that positively correlate with student achievement.

Despite the correlation of students' affective characteristics, previous standardized Mathematics test scores, and grades in past math courses with Advanced Placement scores, there seem to be strategies and techniques that districts, schools, and teachers can employ that offer hope of increasing the likelihood of success for Calculus students (Burney, 2010; Henry, 2008; Peterson, 1989). For example, Escalante utilized increased instructional time and student exposure to great effect in building one of the most impressive Calculus programs in the country in the 1980s (Mathews, 1988). Consequently, there exists both anecdotal evidence and theoretical support for the replication of aspects of Escalante's methodology to narrow or eliminate the gender and ethnic achievement gaps on the Calculus AB exam. Therefore, the focus of this study was whether lengthening of the instructional period promises to reduce the gender and ethnic achievement gaps on the Calculus AB exam.

Chapter 3: Methodology

This chapter contains a detailed explanation of the methodology employed in this study, as well as the design of the study. This study consisted of two components. The first component of the study was a causal-comparative, non-experimental quantitative analysis of extant Calculus AB exam data intended to ascertain whether gender and ethnic achievement gaps, which have been evident in many studies, exist within the state of Michigan, which was the locale of study (Holmes, 2013; Jara, 2013; JBHE Foundation, 2004; Melsom, 2008; Moore & Slate, 2010, 2011; Moore et al., 2012; Morris, 2013; Morris et al., 2015; Wilson, 2013). The second component of the study was a quantitative quasi-experiment to determine whether yearlong daily two-hour blocks of instruction are likely to reduce the size of the twin ethnic and gender achievement gaps on the Calculus AB exam. The organization of this chapter includes 15 sections, beginning with this (a) introduction. Subsequent sections are titled (b) purpose of the study, (c) research questions, (d) hypotheses, (e) research design, (f) selection of participants, (g) instrumentation, (h) data collection, (i) operationalization of variables, (j) data analysis, (k) limitations and delimitations of the research design, (l) internal and external validity, (m) expected findings, (n) ethical issues in the study, and (o) summary.

Purpose of the Study

Previous research has demonstrated that Black and Hispanic students' Advanced Placement exam scores are a function of state residency (Davis, 2012; Koch, 2012; Wilson, 2013). Education is a state responsibility, meaning that each state establishes its own educational policies (Jara, 2013; Morris, 2013). Therefore, researchers have called for investigations of ethnic and gender differences at the state level (Davis, 2012; Jara, 2013; Morris, 2013). To answer the call of those researchers, the purpose of the study was to determine whether changing the length of daily Calculus instruction from one to two hours could affect the ethnic and gender achievement gaps in Calculus AB exam performances in the state of Michigan. As the current study featured two components, an achievement gap analysis and a quasi-experimental component that investigated a possible strategy to reduce the size of the achievement gap, a bifurcated approach was necessary.

The goal of the achievement gap analysis was to determine whether the national Advanced Placement achievement gaps in ethnicity and gender also occur in Michigan on the Calculus AB exam. The findings of the literature review suggest that both gender and ethnic achievement gaps exist throughout the nation, regardless of location (College Board, 2014; Holmes, 2013; Jara, 2013; Melsom, 2008; Moore & Slate, 2010, 2011; Morris et al., 2015). However, the literature does not yet include an analysis of the most recent Calculus AB exam data, nor does the literature provide an analysis of disaggregated Calculus AB exam data for the state of Michigan. Analysis of Calculus AB exam data required the acquisition of Calculus AB data archived by College Board. Consequently, the achievement gap analysis was a dissection of the archived Calculus AB exam data to ascertain whether the historically suggested nationwide gender and ethnic achievement gaps exist within the state of Michigan.

However, determining the existence of Calculus AB exam gender or ethnic achievement gaps within the state of Michigan was the secondary purpose of the study. The primary purpose of this study was to determine whether extended yearlong block periods of Calculus instruction could serve as a strategy to reduce the size of existing gender and ethnic achievement gaps on the Calculus AB exam. The literature suggested that Calculus AB exam scores might increase through yearlong block classes, as previous research demonstrated that students in continuous yearlong block instruction typically outpace their peers in semester classes and increased time on task correlates with higher student achievement (Carroll, 1989; Gullatt, 2006; Hansen et al., 2000; Ottmar et al., 2013; Zelkowski, 2010). Thus, it was necessary to investigate whether gender and ethnic achievement gaps exist in Michigan on the Calculus AB exam. Upon verification of the existence of the gaps, the study proceeded to its primary purpose of investigating whether two-hour blocks of Calculus instruction are a means of reducing the size of gender and ethnic achievement gaps on the Calculus AB exam.

Research Questions

Each of the purposes of this study produced multiple research questions. The first purpose, determining and measuring the existence of Calculus AB exam achievement gaps in Michigan, was responsible for the first three inferential research questions given below. The second purpose, determining whether yearlong two-hour instructional blocks of Calculus can narrow achievement gaps on the Calculus AB exam, produced the final four inferential research questions. Together, the three purposes generated the following seven research questions.

1. What is the difference in the Calculus AB exam performance of Michigan students as a function of gender?

2. What is the difference in the Calculus AB exam performance of Michigan students as a function of ethnicity?

3. What is the difference in the Calculus AB exam performance of Michigan's Hispanic students as a function of gender?

4. What is the difference in Calculus AB exam performance of students as a function of a yearlong two-hour AP Calculus class?

5. What is the difference in Calculus AB exam performance of Hispanic students as a function of a yearlong two-hour AP Calculus class?

6. What is the difference in Calculus AB exam performance of Hispanic male students as a function of a yearlong two-hour AP Calculus class?

7. What is the difference in Calculus AB exam performance of Hispanic female students as a function of a yearlong two-hour AP Calculus class?

Hypotheses

Each of the inferential research questions in the previous section yielded a corresponding null hypothesis. For the sake of clarity, the numeration of the hypotheses in the two subsections below corresponds with that of the research questions. The first subsection in this section contains the seven null hypotheses, while the subsequent subsection lists the corresponding alternate hypotheses.

Null hypotheses.

1. A statistically significant difference in Calculus AB exam performance of Michigan students will not be present as a function of gender.

2. A statistically significant difference in Calculus AB exam performance of Michigan students will not be present as a function of ethnicity.

3. A statistically significant difference in Calculus AB exam performance of Michigan's Hispanic students will not be present as a function of gender.

4. A statistically significant difference in Calculus AB exam performance of students will not be present as a function of a yearlong two-hour AP Calculus class.

5. A statistically significant difference in Calculus AB exam performance of Hispanic students will not be present as a function of a yearlong two-hour AP Calculus class.

6. A statistically significant difference in Calculus AB exam performance of Hispanic male students will not be present as a function of a yearlong two-hour AP Calculus class.

7. A statistically significant difference in Calculus AB exam performance of Hispanic female students will not be present as a function of a yearlong two-hour AP Calculus class.

Alternate hypotheses.

1. A statistically significant difference in Calculus AB exam performance of Michigan students will be present as a function of gender.

2. A statistically significant difference in Calculus AB exam performance of Michigan students will be present as a function of ethnicity.

3. A statistically significant difference in Calculus AB exam performance of Michigan's Hispanic students will be present as a function of gender.

4. A statistically significant difference in Calculus AB exam performance of students will be present as a function of a yearlong two-hour AP Calculus class.

5. A statistically significant difference in Calculus AB exam performance of Hispanic students will be present as a function of a yearlong two-hour AP Calculus class.

6. A statistically significant difference in Calculus AB exam performance of Hispanic male students will be present as a function of a yearlong two-hour AP Calculus class.

7. A statistically significant difference in Calculus AB exam performance of Hispanic female students will be present as a function of a yearlong two-hour AP Calculus class.

Research Design

The research design for this study included two separate components to address the two purposes of the study. This section of this chapter includes a description of the two components of the study. Therefore, the subsequent subsection contains the design of the achievement gap component, while the design of the quasi-experimental component follows in the section after that.

The achievement gap analysis. The first component of the study, the achievement gap analysis, incorporated a quantitative causal-comparative nonexperimental methodology. Causal-comparative research is an ex post facto research design and is appropriate when the researcher is unable to manipulate the independent variable to measure its effect on a dependent variable (Johnson & Christensen, 2013; Silva, 2010). Johnson and Christensen (2013) explain that, "In causal-comparative research, the researcher studies the relationship between one or more categorical independent variables and one or more dependent variables" (p. 44). College Board compiled and collected the Calculus AB Exam data utilized in this study prior to the design of this study, so it was impossible to manipulate the variables within the dataset. Therefore, the appropriate structure for the achievement gap analysis was a quantitative causal-comparative nonexperimental design.

The quasi-experiment. The second component of the research design used a quantitative causal-comparative quasi-experimental model. In this study, the researcher was unable to arrange random assignment of the participants to experimental and control groups, which meant that it was not possible to use an experimental model. Furthermore, the comparison group had already completed their studies in Calculus and their testing using the Calculus AB exam prior to the design of the quasi-experimental component of the study. According to Adams and Lawrence (2015), a quasi-experimental model is a practical alternative when a true experiment is not possible. In light of the debate about the exact variables that contribute to OTL and the difficulty with measuring many of them, the emphasis in this study was on the easily manipulated, and almost universally included, OTL variable of instructional time. Therefore, in the quasi-experimental component of the study, the comparison group received a treatment of yearlong Calculus instruction for one hour each day, while the experiential group received an experiential treatment of yearlong Calculus instruction for two hours each day. The construction of a comparison of the college entrance exam data of the comparison and experiential groups provided a baseline comparison of the two groups. Due to the limited size of the Calculus program at the high school that served as the subject of the study, the comparison group included students from the years 2014 and 2015. Similarly, the experiential group included students from the years 2016 and 2017. Thus, the comparison group attended Calculus classes for one hour per day, whereas the experiential group attend Calculus classes for two hours per day. Following their respective treatments, I conducted a comparison of Calculus AB exam scores to determine whether the experiential treatment produced a difference in the two groups.

The comparison treatment followed a College Board approved AP Calculus AB syllabus and typically consisted of introduction to topics, demonstration of examples, proof of theorems, and recitation to check for understanding. Researchers have found that flipped classrooms have yielded improved results for college level Calculus (Albalawi, 2018; Sahin et al., 2015). Therefore, the first hour of instruction for the experiential group was largely the same as the single hour for the comparison group, as the teacher primarily used that hour for introduction to topics, demonstration of example, proof of theorems, and recitation to check for understanding. However, following the example of the flipped classroom, the teacher mostly used the extra hour of class time with the experiential group to provide students the opportunity to hone their skills in class through regular assignment of classwork, review of students work, and opportunities to correct mistakes and clear up misconceptions.

Selection of Participants

The first component of this study, the achievement gap analysis, used College Board's archived Calculus AB exam data. The subject of the achievement gap analysis was specifically the subgroups of Michigan students, including gender and ethnicity as reported by the individual students themselves, who took the Calculus AB exam in 2002, 2007, 2012, and 2017. In the analysis of said data, the participants were the students who took the Calculus AB exam in each year of the study. The study did not exclude any of the test takers, save those students who failed to provide their gender, ethnicity, or both. So all students who took the Calculus AB exam and identified their gender were included in the gender achievement gap analysis. Likewise, all students who took the Calculus AB exam and identified their ethnicity as Asian, Black, Hispanic, or White were included in the ethnic achievement gap analysis; and, all students who identified both their gender and their ethnicity as Hispanic were included in the Hispanic achievement gap analysis.

The second component of the study, the quasi-experimental component, used the Calculus AB exam data from a Michigan high school with a sizable Hispanic population. The participants in the comparison group were Calculus AB exam takers from consecutive years, 2014 and 2015, where Calculus instruction lasted one hour each day. Therefore, the comparison group included 56 students, 41 students who took the 2014 Calculus AB exam and another 15 students who took the 2015 Calculus AB exam. Correspondingly, the participants in the experiential group were Calculus AB exam test takers from consecutive years, 2016 and 2017, where Calculus instruction lasted two hours each day. Hence, the experiential group included 56 students, 25 students who took the 2016 Calculus AB exam and 31 more students who took the 2017 Calculus AB exam. Once again, the entire population within both the comparison group and the experiential group was included in the study. Notably, all participants in the quasi-experimental component attended a Title I school.

Instrumentation

The first component of this study, the achievement gap analysis, used archival data from the Calculus AB exam downloaded from the College Board website. To ascertain the existence of achievement gaps between genders and ethnicities on the Calculus AB exam in the state of Michigan, this study included data from 2002, 2007, 2012, and 2017. Therefore, the instrument for data collection for the achievement gap analysis was the Calculus AB exam for the aforementioned years.

The second component of the study, the quasi- experimental portion, used scores from college entrance examinations to compare the comparison and experiential groups. However, for the comparison group, only ACT scores were available, while for the experiential group, only SAT scores were available. There are differences in the way that scores are reported for the two tests, as ACT scaled scores for mathematics range from 1 to 36, while SAT scores for mathematics range from 200 to 800 (ACT, 2014; College Board, 2015d). To compare the scores, College Board recommends the usage of concordance tables instead of direct comparisons of percentile ranks (College Board, 2015c).

The quasi-experimental component of the study also used school level Calculus AB exam data that was not publicly accessible to compare the achievement of the groups. Specifically, the quasi-experiment used Calculus AB exam data gathered from the comparison group for the years

2014 and 2015 and from the experiential group for the years 2016 and 2017. Hence, the data-gathering instruments for the quasi-experimental component of the study were the ACT for 2014 and 2015; the SAT for 2016 and 2017; and, the Calculus AB exam for the years 2014 to 2017.

Because the achievement gap analysis and the quasi-experiment both utilized the Calculus AB exam as the instrument for data collection, the remainder of the instrumentation section consists of three subsections pertaining to the primary instrument of this study, which is the Calculus AB exam. The first subsection describes the Calculus AB exam, the second subsection addresses the reliability of the instrument, and the third subsection will discuss the validity of the instrument. After this section, the next section contains a description of the data collection in this study.

Description of the instrument. The format of the Calculus AB exam remained largely the same for the duration of the study. The Calculus AB exam consists of two sections, a multiple-choice section and a free-response section (College Board, 2012, 2016a). The multiple-choice section of the Calculus AB exam contains two parts (College Board, 2012, 2016a). Part A of the multiple-choice section is a non-calculator subsection, while Part B is a subsection requiring a calculator (College Board, 2012, 2016a). From 2002 to 2016, Part A of the multiple-choice section contained 28 questions, while Part B consisted of 17 questions (College Board, 2012). Although the structure changed slightly for the 2017 exam, the number of multiple-choice questions remained unchanged. For the 2017 administration of the Calculus AB exam, the number of questions in Part A of the multiple-choice section was increased to 30, while the number of questions in Part B was decreased to 15 (College Board, 2016a).

The second section of the Calculus AB exam is a free response section, which also consists of two subsections (College Board, 2012, 2016a). Currently, Part A of the free response section incorporates two free response questions requiring a graphing calculator and Part B includes four free response questions without a calculator (College Board, 2012, 2016a). College Board (2017a) structured the Calculus AB exam in this same manner for each year from 2011 to 2017. However, from 2002 to 2010, College Board (2017a) divided the free response questions evenly between Part A, which required a calculator, and Part B, which was a non-calculator section.

Despite the changes in the Calculus AB exam, the scoring of the exam has remained the same. This is possible because the number of multiple-choice questions remained unchanged at 45, while the number of free response questions was six for each year of the study. Regardless of the composition of the multiple choice and free response sections, each section contributed half of the points to the overall score (College Board, 2012, 2016a). College Board (2012, 2016a) converts students' raw scores to scaled scores ranging from 1 to 5. Thus, students who get a scaled score of 1 receive no recommendation and students who score a 2 are labeled possibly qualified by College Board (2012, 2016a). Similarly, students who score a 3 are given a recommendation of qualified, students earning a scaled score of 4 earn a recommendation of well qualified, and students receiving a score of 5 are awarded a recommendation of extremely well qualified by College Board (2012, 2016a).

Instrument reliability. In educational research, the reliability of the data-gathering instrument is critical (Adams & Lawrence, 2015; Johnson & Christensen, 2013). The reliability of a study stands on the ability of the instrument to produce consistent results (Adams & Lawrence, 2015). In the case of the Calculus AB exam, College Board issues a different form of the exam each year. Therefore, it is crucial to establish alternate form reliability to ensure that the various forms of the exam used throughout a study produce consistent results (Adams &

Lawrence, 2015). In short, the Calculus AB exam should produce consistent results over time so that equivalent levels of proficiency result in the same score (Koch, 2012; Wilson, 2013).

To determine the reliability of the Calculus AB exam, I sent an external request to College Board after the completion of the 2017 Advanced Placement testing cycle. Typically, College Board provides coefficient alpha values for Advanced Placement exams (Davis, 2012; Koch, 2012). After making similar requests, Davis (2012) and Koch (2012) determined that the Advanced Placement exams have a coefficient alpha in excess of .70, which is the minimum that Johnson and Christensen (2013) recommend for establishing reliability. College Board officials provided the same measures to the request that I made for the purposes of this study as they did for Davis (2012) and Koch (2012). Coefficient alpha measures were in excess of .70 for all measures of internal consistency for all seven years of data analyzed within the course of the study, while all score boundaries had coefficient alpha measures exceeding .90 for all seven years of the study. Consequently, the Calculus AB exam was a suitably reliable instrument for comparing student achievement across an extended amount of time.

Instrument validity. Since an instrument can be reliable without being valid, it is also important to establish the validity of the instrument (Adams & Lawrence, 2015; Johnson & Christensen, 2013). Validity refers to the accuracy of the measures delivered by the instrument (Adams & Lawrence, 2015; Johnson & Christensen, 2013). An important category of validity is construct validity, which refers to whether a given concept behaves like a variable (Adams & Lawrence, 2015). Construct validity is frequently assessed based upon content and criterion (Adams & Lawrence, 2015). The analysis of validity for the Calculus AB exam will examine each of these components of construct validity separately.

The Calculus AB exam measures the degree to which high school students have learned the skills that professors teach in a typical college Calculus course (College Board, 2016a). Keng and Dodd (2008) found that the students who earned college credit due to their Calculus AB exam scores earned higher grades in the sequent course than their peers who were unable to earn passing scores on the Calculus AB exam. Morgan and Klaric (2007) found that students who took Advanced Placement Calculus and were awarded credit for their Calculus AB exam score completed more courses in a related field that their peers who did not take the Advanced Placement course. Consequently, the Calculus AB exam owns a history of correlation between test results and college readiness, which suggests content validity of the instrument.

Criterion validity requires a, "Positive correlation between scale scores and a behavioral measure" (Adams & Lawrence, 2015, p. 94). The Calculus AB exam exhibits this phenomenon as an increase in scale scores on the exam positively correlates with subject GPA (Keng & Dodd, 2008). Similarly, Morgan and Klaric (2007) reported that increases in the Calculus AB exam score correlated positively with both differences in course performance and in SAT-adjusted course performance. Therefore, there is historical evidence that the Calculus AB exam has criterion validity.

Data Collection

After the Institutional Review Board at Concordia University-Portland gave permission to proceed, I downloaded the requisite data files from the College Board website. For the first component of the study, I obtained disaggregated data from the Calculus AB exam from the College Board website in Excel files. To determine the existence of gender and ethnic achievement gaps in the state of Michigan, I downloaded separate statewide data Excel files for the Calculus AB exam for the years 2002, 2007, 2012, and 2017. College Board disaggregates statewide data by both gender and ethnicity.

For the quasi-experimental component, I downloaded school level data from College Board's website for the high school that served as the subject of the study. This required me to download a separate school level data file for each year of the quasi-experimental portion of the study. Therefore, I downloaded separate data files for each of the comparison years, which were 2014 and 2015, and two additional files for the experiential years, which were 2016 and 2017. Unlike with the state level data files, College Board did not provide disaggregated school level Excel data files. Thus, I disaggregated the files containing school level data manually.

Operationalization of Variables

Some authors and researchers note that occasionally some researchers utilize means for ordinal data when the calculation yields meaningful results, but interpretation of those calculations warrants caution (Remler & Van Ryzin, 2014; Rubin, 2014). In contrast, many statisticians and researchers believe that serious studies should not include the calculation of means for ordinal data; and instead, they suggest focusing on the median as the measure of central tendency (Adams & Lawrence, 2015; Gravetter & Wallnau, 2009). Despite the controversy, each year College Board disaggregates Calculus AB exam data, which is ordinal data, by both gender and ethnicity and subsequently reports means for each subgroup. However, in this study I made a concerted effort to avoid the usage of the mean as a measure of central tendency and instead utilized the median and distribution of the variable.

Gender and ethnicity are clearly nominal categorical data, but it is less obvious that Calculus AB exam scores are also categorical variables. The scores reported to students are numerical values that represent ordinal categories that communicate the recommendation of

College Board (College Board, 2012, 2016a). Therefore, Advanced Placement scores are essentially a Likert scale that translates the five ordinal categories of no recommendation, possibly qualified, qualified, well qualified, and extremely well qualified into integral values ranging from 1 to 5 (College Board, 2012, 2016a).

Each component of the study involved the Calculus AB exam, but the dependent variables for the two components were different. In the first component of the study, the achievement gap analysis, the independent variables are gender and ethnicity, meanwhile Calculus AB exam scores serve as the dependent variable. Similarly, in the quasi-experimental component, the independent, or manipulated, variable is the duration of each instructional period, while Calculus AB exam scores serve as the dependent variable.

Data Analysis Procedures

College Board (2016a) annually reports Calculus AB exam score distributions with integral scores ranging from 1 to 5. Nevertheless, some researchers reason that since scores of both 1 and 2 receive no college credit, the difference between the two scores is negligible (Davis, 2012; Holmes, 2013; Jara, 2013; Morris, 2013). In a similar fashion, Wilson (2013) chose to lump Advanced Placement exam scores into two categories, passing and non-passing. Wilson (2013) grouped scores of 1 to 3 into the non-passing category and scores of 4 and 5 into the passing category. However, some schools, like Central Michigan University (2015), Ferris State University (2016), University of Michigan-Dearborn (2017), and Wayne State University (2015), choose to award credit for a Calculus AB score of 3, while others, such as Michigan State University (2017) and Kalamazoo College (2017) award credit for a 4 or better, and still others, like the University of Michigan (2017), only award credit for a 5. In contrast, Buford (2012) and Koch (2012) recognized the value of the various scores and chose to use College Board's

(2016a) five-point scale. Therefore, Davis (2012), Holmes (2013), and Morris (2013) all chose to conduct Chi-square calculations with four categories of scores, Wilson (2013) operated with two categories, and Koch (2012) opted for five categories. Following in the footsteps of Koch (2012) and recognizing the potential value of each score in measuring incremental improvement in an Advanced Placement program, the analysis in this study consists of five categories for each Chi-square analysis and Fisher-Freeman-Halton exact probability test. Since Calculus AB exam scores, as well as both gender and ethnicity, are categorical data, Pearson Chi-square tests were to be conducted to determine whether the dependent variable, Calculus AB exam score distributions, were statistically significantly different based on the independent variable, either gender or ethnicity.

The intended purpose of the first component of this study, the achievement gap analysis, was to ascertain whether gender and ethnic achievement gaps on the Calculus AB exam exist in the state of Michigan. The evaluation of these gaps required acquisition of archived Calculus AB exam data from College Board. The next paragraph contains a description of the data analysis procedures for the first three research questions, while the two paragraphs after that contain a description of the calculation of effect size and its interpretation.

For the first research question, 2 (gender) x 5 (exam score) Chi-square procedures were conducted to ascertain whether Calculus AB exam scores were independent from gender. For the second research question, 4 (ethnicity) x 5 (exam score) Chi-square procedures were conducted to ascertain whether Calculus AB exam scores were independent from ethnicity. Four categories of ethnicity were included in this study: Asian, Black, Hispanic, and White. For older data sets, the Hispanic category consisted of a merger of three Latino groups, namely Chicano / Mexican American, Puerto Rican, and Other Latino. For the third research question, 2 (gender) x 5 (exam score) Chi-square procedures were conducted to ascertain whether Calculus AB exam scores were independent from gender for Hispanic students. Once again, the Hispanic category for older data sets consisted of a merger of three Latino groups, namely Chicano / Mexican American, Puerto Rican, and Other Latino.

The next step was to determine the statistical significance of each of the Chi-square tests. Statistical significance is a measure of the likelihood that a given result did not occur by chance (Adams & Lawrence, 2015; Johnson & Christensen, 2013). The statistical convention has become to use a significance level of .05 (Adams & Lawrence, 2015; Johnson & Christensen, 2013). This means that the researcher should reject the null hypothesis if the probability of the result is less than 5% (Adams & Lawrence, 2015). The size of the sample is an important variable in determining the practical significance of a statistically significant finding (Adams & Lawrence, 2015; Johnson & Christensen, 2013). Thus, each researcher should calculate effect size to indicate the practical significance of a result (American Psychological Association, 2010).

In the Chi-square computations, the contingency tables were not square, which meant that the correct measure of effect size was Cramer's V, squared (Adams & Lawrence, 2015). Cohen outlined a criteria for interpreting effect size, where .10 represents small effect, .30 represents medium effect, and .50 represents large effect size (Bosco, Aguinis, Singh, Field, & Pierce, 2015; Gignac & Szodorai, 2016). In contrast, Gignac and Szodorai (2016) recommend using .10 for small effect, .20 for medium effect, and .30 for large effect. However, many researchers accept Cohen's criteria as the convention to determine effect size and continue to use it. (Bosco et al. 2015; Gignac & Szodorai, 2016).

The intended purpose of the second component of this study, the quasi-experimental component, was to determine whether an experiential treatment consisting of yearlong two-hour blocks of Calculus instruction could reduce the size of the achievement gaps. The data for this component of the study emanated from two independent groups, namely a comparison group and an experiential group. The quasi-experimental component consisted of two sets of data analysis, a comparison of college entrance examination scores to determine whether the two groups were statistically similar and a comparison of Calculus AB exam distributions to determine whether the treatment resulted in a statistically significant difference.

First, I translated the ACT scores for experiential group were into SAT scores using the concordance table provided by College Board (2015c, 2016b). Subsequently, I grouped the scores for both the comparison group and the experiential group into five intervals: 301 to 400, 401 to 500, 501 to 600, 601 to 700, and 701 to 800. Next, score distributions for the comparison group and the experiential group were to be compared with a 2 (comparison and experiential groups) x 5 (groups of SAT scores) Chi-square test. Due to the high level of expected cell frequencies that were less than five, Fisher-Freeman-Halton exact probability tests replaced Chi-square analysis (Fagerland, Lydersen, & Laake, 2017). Following statistical convention, I concluded that any difference between the distributions of scores for the comparison and experiential groups was statistically significance the .05 level (Adams & Lawrence, 2015). Then, to analyze the Calculus AB exam data for the two groups, a Chi-square test, which is the appropriate computational tool to measure whether the experiential treatment produced a statistically significant difference, was to be conducted (Adams & Lawrence, 2015).

For the fourth research question, 2 (comparison and experiential groups) x 5 (exam score) Chi-square procedures were to be conducted to ascertain whether there was a statistically significant difference in Calculus AB exam performance of students as a result of taking a two hour per day Calculus class instead of a one hour class. However, the data produced a high level of expected cell frequencies that were less than five, meaning that Chi-square analysis was inappropriate for the data (Fagerland et al., 2017). Fagerland et al. (2017) indicate that the correct statistical analysis tool for this data set is Fisher-Freeman-Halton exact probability test. Consequently, Fisher-Freeman-Halton exact probability tests replaced Chi-square calculations as the analytical tool of choice for the fourth research question.

For the fifth research question, 2 (comparison and experiential groups) x 5 (exam score) Chi-square procedures were to be conducted to ascertain whether there was a statistically significant difference in Calculus AB exam performance of Hispanic students as a result of taking a two hour per day Calculus class instead of a one hour class. However, the data produced a high level of expected cell frequencies that were less than five, meaning that Chi-square analysis was inappropriate for the data (Fagerland et al., 2017). Fagerland et al. (2017) indicate that the correct statistical analysis tool for this data set is Fisher-Freeman-Halton exact probability test. Consequently, Fisher-Freeman-Halton exact probability tests replaced Chi-square calculations as the analytical tool of choice for the fifth research question.

For the sixth research question, 2 (comparison and experiential groups) x 5 (exam score) Chi-square procedures were to be conducted to ascertain whether there was a statistically significant difference in Hispanic male student scores on the Calculus AB exam as a result of taking a two hour per day Calculus class instead of a one hour class. However, the data produced a high level of expected cell frequencies that were less than five, meaning that Chi-square analysis was inappropriate for the data (Fagerland et al., 2017). Fagerland et al. (2017) indicate that the correct statistical analysis tool for this data set is Fisher-Freeman-Halton exact probability test. Consequently, Fisher-Freeman-Halton exact probability tests replaced Chi-square calculations as the analytical tool of choice for the sixth research question.

For the seventh research question, 2 (comparison and experiential groups) x 5 (exam score) Chi-square procedures were to be conducted to ascertain whether there was a significant difference in Hispanic female student scores on the Calculus AB exam as a result of taking a two hour per day Calculus class instead of a one hour class. However, the data produced a high level of expected cell frequencies that were less than five, meaning that Chi-square analysis was inappropriate for the data (Fagerland et al., 2017). Fagerland et al. (2017) indicate that the correct statistical analysis tool for this data set is Fisher-Freeman-Halton exact probability test. Consequently, Fisher-Freeman-Halton exact probability tests replaced Chi-square calculations as the analytical tool of choice for the seventh research question.

Once again, I followed statistical convention and a significance level of .05 to determine whether observed differences were statistically significant. Next, I computed the effect size of each measure. Since none of the contingency tables were square for each of the last four research questions, the correct measure of effect size was Cramer's V, squared (Adams & Lawrence, 2015). Once again, Cohen's convention was used to determine the practical significance of any statistically significant finding (Bosco et al. 2015; Gignac & Szodorai, 2016).

Limitations and Delimitations of the Research Design

This section lists the limitations and delimitations of the research design of this study. The limitations included assumptions connected to the instrumentation, unmeasured cultural effects, and quasi-experimental assumptions. After discussion of each of the aforementioned limitations, the delimitations of the study are given.

Limitations. The limitations on this study include two assumptions pertaining to instrumentation. The first assumption was that the Calculus AB exam scores reported by College Board were both reliable and valid. Therefore, the next section of this chapter contains discussion pertaining to the reliability and validity of the instrument. The second assumption was that the minor changes that College Board made in the format of the Calculus AB exam had no effect on student scores.

Another limitation of this study was the inability of the study to measure cultural effects realized by the students themselves. One example of this is the potential for societal norms to influence both the participation in an Advanced Placement Calculus class and the likelihood that students will opt to take the Calculus AB exam. Another example is parental support toward the goal of passing the Calculus AB exam. Therefore, while the study may find differences in exam scores between the two genders and the various ethnicities, it is unable to clarify the reason for those differences.

The quasi-experimental component of this study had additional limitations. At the inception of the quasi-experiment, ACT scores were supposed to be used to determine if the comparison group and the experiential group were statistically equivalent. At the inception of the quasi-experiment, the state of Michigan required all students to take the ACT during their junior year (Goodman, 2016). Therefore, I chose to utilize the ACT to determine statistical equivalence between the comparison group and the experiential group. However, after the first two years of the quasi-experiment, the Michigan Legislature elected to replace the ACT test with the SAT test (Woods, 2015). Dorans (2004) calculated that the ACT and SAT math sections had a correlation coefficient of .89; however, there have been changes to both instruments since his evaluation. Therefore, a limitation of this study is the accuracy of the concordance tables that convert scores between the two college entrance examinations. Consequently, this resulted in a limitation based upon any demonstrated statistical equivalence between the comparison and experiential groups due to similarity in college entrance examination scores and completion of the Algebra 1, Geometry, Algebra 2, and Precalculus course sequence.

Another limitation pertaining to the quasi-experimental component was the inability of the research design to measure the quality of instruction was the same for both the comparison and experiential groups. Since the same teacher taught both groups of students, but in pairs of sequent years, an assumption that was included was that the teacher did not improve his instructional techniques or pedagogy during the course of the four years that comprised the quasi-experiment. Similarly, another assumption that was included in the study was that changes in content had no effect on Calculus AB exam scores. In summation, for the purposes of this study, I assumed that differences in Calculus AB exam scores between the comparison and experiential groups were attributable to the increased amount of time for instruction.

Delimitations. The principal delimitations of this study were the restrictions to only Calculus AB exam scores within the state of Michigan. The achievement gap analysis included a delimitation to students who reported their gender for the research questions that addressed gender and to students who reported their ethnicity for the questions that addressed ethnicity. Moreover, the gap analysis component of the study included a delimitation to students who took the Calculus AB exam in the years 2002, 2007, 2012, and 2017. The quasi-experimental component included a delimitation to a single high school in Michigan and the students who took the Calculus AB exam between the years of 2014 and 2017. I chose the schedule structures out of convenience and this resulted in a delimitation of the comparison group to one hour of Calculus instruction, while the experiential group received two hours of daily Calculus instruction. Additionally, the quasi-experimental component included an investigation of the impact of OTL on Calculus AB exam achievement, but the quasi-experiment included a delimitation the variable of instructional time.

Internal and External Validity

Examination of the validity of a study provides insight into how researchers can replicate aspects of a study and identify potential methods to improve future research (Onwuegbuzie, 2000). Nine threats to internal validity are "history, maturation, testing, instrumentation, selection, mortality, diffusion of treatment and compensatory equalization, rivalry and demoralization" (Drost, 2011, p. 115). Threats to the internal validity of this particular study included instrumentation, differential selection of participants, and experimental mortality. Two threats to external validity were population validity and ecological validity.

While there is sufficient evidence that the Calculus AB exam is a valid instrument for measuring whether high school students have attained a college equivalent level of Calculus proficiency, there was a threat to the validity of the study attached to the instrumentation. One threat was due to the changes in the structure of the Calculus AB exam described in the instrumentation section (Drost, 2011; Onwuegbuzie, 2000). The validity threat due to instrumentation potentially affected both the achievement gap analysis and the quasi-experimental component. An additional threat to the validity of the quasi-experimental section was due to the usage of different instruments to compare the statistical equivalence of the comparison and experiential groups.

Selection bias pertained to both components of the study. The selection of participants in this study was not random, which may have inadvertently resulted in differential selection of participants (Drost, 2011; Onwuegbuzie, 2000). Moreover, students self-selected whether they would take the Calculus AB exam, which contributes bias to the selection process. There was an additional threat in the quasi-experimental component because the comparison group and experiential groups took the Calculus AB exam in different years. Additionally, I chose the school that was the subject of the quasi-experimental component out of convenience.

Students' self-selection of taking the Calculus AB exam also represented experimental mortality (Drost, 2011; Onwuegbuzie, 2000). Experimental mortality occurs when participants choose not to complete the experiment (Drost, 2011; Onwuegbuzie, 2000). Experimental mortality affected both components of the study.

Onwuegbuzie (2000) explained that population validity consisted of the ability to apply conclusions from a study in other settings. The achievement gap analysis analyzed data from the state of Michigan, so findings from this state may not apply to other locales. Similarly, the quasi-experimental component of this study took place in a high school in Michigan and conclusions drawn from this research may not transfer to other settings.

Expected Findings

Multiple studies report that the Calculus AB exam exhibits gender differences with males outscoring females (Moore et al., 2012; Morris, 2013). Thus, the literature suggested that the achievement gap analysis for the Calculus AB exam in Michigan would uncover a gender gap that demonstrated that male students achieve higher scores than female students. In the analysis of ethnic exam data, multiple studies show that both Asian and White students outscore both Hispanic and Black students on the Calculus AB exam (Moore & Slate, 2008; Holmes, 2013; Jara, 2013). Hence, the literature suggested that both White and Asian students would achieve higher scores on the Calculus AB exam than their Black and Hispanic peers.

OTL research suggested that increased time of instruction should result in increased achievement, especially since the instruction was to be yearlong (Gullatt, 2006; Hansen et al., 2000; Lightner, 2009; Smith & Camara, 1998). Researchers' conclusions have suggested that students score higher on the Calculus AB exam when yearlong block instruction replaced a single period of instruction that lasted an hour or less (Gullatt, 2006; Hansen et al., 2000; Keen, 1996; Lightner, 2009). Thus, in the quasi-experimental component, there was an expectation that the experiential treatment would produce higher scores on the Calculus AB exam than the comparison treatment.

Ethical Issues in the Study

This study consisted of two components, an achievement gap analysis and a quasi-experiment. The achievement gap analysis required the download and analysis of Calculus AB exam data. As the names of the participants are not identifiable and there is no risk of damage to the participants, the achievement gap analysis was a category that was exempt (Johnson & Christensen, 2013). Thus, consent forms were not required for the first component. In contrast, the quasi-experimental component assigned participants to take either one or two hours of Calculus instruction daily. However, school officials made the decision to schedule the courses, which resulted in typical education practices within the context of a typical school setting, which was also an exempt category (Johnson & Christensen, 2013). Although both components of the study fell into exempt categories, the Institutional Review Board still had to approve the study, as it does for all research involving human subjects (Johnson & Christensen, 2013).

Summary

One purpose of this study was to investigate ethnic and gender achievement gaps pertaining to the Calculus AB exam in the state of Michigan. To address this purpose, the study included an achievement gap analysis. The achievement gap analysis was a causal comparative quantitative data analysis to determine whether there were gender and ethnic achievement gaps on the Calculus AB exam in Michigan. The achievement gap analysis featured Calculus AB exam data from 2002, 2007, 2012, and 2017. The achievement gap analysis used Chi-square procedures to identify discrepancies between distributions. Effect sizes were calculated for statistically significant results, which were subsequently interpreted using Cohen's criteria (Bosco et al. 2015; Gignac & Szodorai, 2016).

The second component of this study was a quasi-experimental component that dissected the Calculus AB exam achievement data associated with a shift from one hour of instruction to two hours of daily Calculus instruction. The quasi-experimental component extended for a period of four years, as the comparison group tested in 2014 and 2015, while the experiential group tested in 2016 and 2017. To determine whether the two groups were statistically equivalent prior to treatment, I compared College entrance exam data for the comparison and experiential groups. The comparison group received a single hour of daily Calculus instruction, while the experiential group received two hours of daily Calculus instruction. To determine whether there was a statistically significant difference between the two treatments, I conducted Chi-square analysis. Where necessary, I substituted Fisher-Freeman-Halton exact probability tests for Chi-square procedures. When a difference was statistically significant, then I determined effect size by calculating Cramer's V, squared. Finally, the effect size was interpreted using Cohen's criteria (Bosco et al. 2015; Gignac & Szodorai, 2016).

Chapter 4: Data Analysis and Results

The purpose of the study was to examine ethnic and gender differences in Advanced Placement Calculus AB exam performances in the state of Michigan. Two components were included in the study: a statewide achievement gap analysis and a single school quasi-experimental component to test whether increasing the length of Advanced Placement Calculus might serve as a means of shrinking or eliminating achievement gaps. The achievement gap analysis consisted of analysis of pre-existing test data to determine whether there were achievement gaps in Calculus AB exam performance in the state of Michigan. The quasi-experimental component was an investigation into whether increasing the length of Advanced Placement Calculus from one hour to two hours daily would produce differences in achievement that might serve as a possible remedy to the problem of achievement gaps.

The achievement gap analysis included Calculus AB exam results from the state of Michigan for the years 2002, 2007, 2012, and 2017. The achievement gap analysis consisted of three separate comparisons. First, I analyzed the data from the four years to determine whether Michigan had a gender achievement gap on the Calculus AB exam. Second, I analyzed the data from the four years to determine whether Michigan had an ethnic achievement gap on the Calculus AB exam. Third, I examined the data from the four years was to determine whether Michigan had a Hispanic gender achievement gap on the Calculus AB exam. Each of the three achievement gap analyses consisted of Chi-square analysis to determine whether any differences between groups was statistically significant.

The quasi-experimental component consisted of a comparison group that received one hour of daily Calculus instruction and an experiential group that received two hours of daily Calculus instruction. The comparison group consisted of participants who took Advanced

Placement Calculus in either of the school years of 2013–2014 or 2014–2015, while the experiential group took the class in either 2015–2016 or 2016–2017. To determine whether the experiential treatment resulted in different score distributions for all participants, Hispanic participants, Hispanic male participants, and Hispanic female participants, I compared the score distributions for the comparison and experiential groups. I planned to use Chi-square analysis for each of the four comparisons, but due to a high level of expected cell frequencies that were less than five, it was necessary to replace Chi-square analysis with Fisher-Freeman-Halton exact probability tests (Fagerland et al., 2017).

The organization of this chapter includes five sections, beginning with this introduction. The chapter continues with sections entitled description of the sample, a summary of the results, and detailed analysis. The chapter concludes with a summary of the chapter.

Description of the Sample

The achievement gap analysis was a census of the students from the state of Michigan who took the Calculus AB exam in each of the years 2002, 2007, 2012, and 2017. All of the participants were students who elected to enroll in an Advanced Placement Calculus course at their high school, and subsequently chose to take the Calculus AB exam; consequently, the selection of the participants was not a random event. Similarly, the subpopulations for each year were students who elected to mark their ethnicity and gender on their answer documents. Not all of the student participants marked their gender, so I excluded participants who declined to supply their gender from the gender achievement gap analysis. Similarly, I excluded participants who declined to supply their ethnicity from the ethnic achievement gap analysis. Likewise, only students who both indicated that their ethnicity was Hispanic and also supplied their gender were included in the Hispanic gender achievement gap analysis. The next paragraph will describe the sample for the gender achievement gap analysis, the subsequent paragraph will describe the sample for the ethnic achievement gap analysis, and the paragraph after that will describe the sample for the Hispanic gender gap analysis.

Table 3

Gender Gap Participant Demographics

Year	Number of Females	Number of Males	Percent of Females	Percent of Males
2002	2,215	2,629	45.7%	54.3%
2007	3,262	3,624	47.4%	52.6%
2012	3,610	4,152	46.5%	53.5%
2017	4,022	4,400	47.8%	53.2%

Presented in Table 3 is a demographic summary of the participants in the gender gap analysis. For the year 2002, there were 4,844 total participants and there were 414 more males than females, which resulted in a difference of 8.6 percentage points. For the year of 2007, there were 6,886 total participants and there were 362 more males than females, which resulted in a difference of 5.2 percentage points. For the year of 2012, there were 7,762 total participants and there were 542 more males than females, which resulted in a difference of 7.0 percentage points. For the year of 2017, there were 8,422 and there were 378 more males than females, which resulted in a difference of 4.4 percentage points. There was a monotonic increase in participation for both genders, but male participants outnumbered female participants in all four years of the study, with the disparity ranging from a difference of 4.4 to 8.6 percentage points. Notably, the widest margin occurred in 2002, the earliest year of the study, while the narrowest margin occurred in 2017, the most recent year of the study.

Table 4

Ethnic Gap Demographics

Year	Asian	Black	Hispanic	White	Total
2002	409	134	86	3,952	4,581
2007	583	259	136	5,534	6,512
2012	634	249	210	6,218	7,311
2017	898	327	390	6,402	8,017

Table 5

Ethnic Gap Percentages

Year	Asian	Black	Hispanic	White
2002	8.0%	2.9%	1.9%	86.3%
2007	9.0%	4.0%	2.1%	85.0%
2012	8.7%	3.4%	2.9%	85.0%
2017	11.2%	4.1%	4.9%	79.9%

Presented in Table 4 is a demographic summary of the participants in the ethnic gap analysis. The number of White, Asian, and Hispanic participants increased with each successive year. The number of Black participants dipped from 2007 to 2012. However, as can be seen by the percentages of participants by ethnicity for each year shown in Table 5, the only ethnicity that saw strictly monotonic increases in the percentage of total participants throughout the course of the four analyzed years was Hispanic. For 2002, the overwhelming majority of participants indicated that their ethnicity was White. Less than 2% of participants indicated that their ethnicity was Hispanic. From 2002 to 2007, there was a slight decline in the percentage of White participants, which accompanied an even smaller increase in the percentage of Hispanic participants. From 2007 to 2012, the percentage of participants that were White remained the same, while there was almost a full percentage point increase in Hispanic participation. From

2012 to 2017, White participation declined by just over five percentage points, while Hispanic participation increased by two full percentage points.

Table 6

Hispanic Gender Gap Participant Demographics

Year	Number of Females	Number of Males	Percent of Females	Percent of Males
2002	37	49	43.0%	57.0%
2007	54	83	39.4%	60.6%
2012	99	111	47.1%	52.9%
2017	182	208	46.7%	53.3%

Presented in Table 6 is a demographic summary of the participants in the Hispanic gender gap analysis. For the year 2002, there were 86 total participants in the Hispanic gender gap analysis. There were 12 more males than females, which resulted in a 14-percentage point difference in participation. For the year of 2007, there were 137 total participants in the Hispanic gender gap analysis. There were 29 more males than females, which resulted in a 21.2 percentage point difference in participation. For the year of 2012, there were 210 total participants in the Hispanic gender gap analysis. There were 22 more males than females, as the percentage point difference narrowed to only 5.8. For the year of 2017, there were 390 total participants in the Hispanic gender gap analysis. There were 46 more male participants than female participants, which resulted in a 6.6 percentage point difference. Hence, there was an increase in the number of both male and female participants each year of the study, although male participants outnumbered female participants in all four years of the study, with the disparity ranging from a difference of 5.8 to 14.0 percentage points.

The quasi-experimental component of the study featured a census of the students who took the Calculus AB exam at a high school in Michigan whose population was over 90%

Hispanic. The comparison group included participants who enrolled in Advanced Placement Calculus and took the AB exam in either 2014 or 2015. The comparison group consisted of 56 participants, 41 of whom who took the 2014 Calculus AB exam and another 15 participants who took the 2015 Calculus AB exam. The comparison group contained 31 females and 25 males; therefore, the comparison group was 55.4% female and 44.6% male. The comparison group contained 3 students who were White, 0 students who were Black, 1 students who was Asian, and 52 students who were Hispanic. Consequently, the comparison group was 5.4% White, 0.0% Black, 1.8% Asian, and 92.9% Hispanic.

The experiential group included participants who enrolled in Advanced Placement Calculus and took the Calculus AB exam in either 2016 or 2017. The experiential group contained 39 females and 17 males; therefore, the comparison group was 69.6% females and 30.4% male. The experiential group contained 2 students who were White, 4 students who were Black, 0 students who were Asian, and 50 students who were Hispanic. Consequently, the comparison group was 3.6% White, 7.1% Black, 0.0% Asian, and 89.3% Hispanic.

Summary of the Results

The first component of the study consisted of an achievement gap analysis. In the achievement gap analysis, I analyzed data from the years 2002, 2007, 2012, and 2017 to determine whether there were achievement gaps in the state of Michigan on the Calculus AB exam. The achievement gap analysis included three perspectives: gender, ethnicity, and gender in the Hispanic population. The next three paragraphs document the results of the achievement gap analysis, after which is the presentation of the results of the quasi-experimental component.

The results of the gender achievement gap analysis indicated that a statistically significant difference between genders in Calculus AB exam scores existed in each of the four years that was tested. The effect size for the gender achievement gap analysis ranged from trivial to small. The results of the ethnic achievement gap analysis indicated that a statistically significant difference between ethnicities in Calculus AB exam scores existed in each of the four years that was tested. The effect size for the ethnic achievement gap analysis ranged from trivial to small. The results of the Hispanic gender achievement gap analysis indicated that a statistically significant difference existed between genders in Calculus AB exam scores for 2017, but not for 2002, 2007, and 2012; although, the effect size for the difference in 2017 was small.

The second component of the study was a quasi-experiment to determine whether increasing the length of the instructional period might affect the size of the gender and ethnic achievement gaps. This component was an investigation into an experiential treatment of two hours of Calculus instruction would produce statistically significant difference in Calculus AB scores than a comparison treatment of a single hour of instruction. Specifically, the researcher compared the scores of all students, Hispanic students, Hispanic males, and Hispanic females.

For the quasi-experimental component, a statistically significant difference existed between participants in the comparison and the experiential groups in Calculus AB exam scores. For Hispanic students, a statistically significant difference existed between participants in the comparison and the experiential groups in Calculus AB exam scores. Additionally, for both Hispanic male students and Hispanic female students there was a statistically significant difference in Calculus AB exam scores between participants in the comparison and experiential groups. Effect sizes for the statistically significant differences found in the quasi-experimental component ranged from medium to large.

In general, the results from the achievement gap analysis led to the conclusion that there are both gender and ethnic achievement gaps on the Calculus AB exam in Michigan.

Specifically, male students outscored female students in the gender gap analysis and Asian and White students outscored Hispanic and Black students in the ethnic gap analysis, while Hispanic males outperformed Hispanic females only in the most recent year of the study. Regarding the quasi-experimental component, the results indicated that the experiential treatment generated a statistically significant difference in Calculus AB exam scores for every category in the investigation.

Detailed Analysis

This study consisted of two components, an achievement gap analysis and a quasi-experiment. The first three research questions emanated from the achievement gap analysis, while the final four research questions stemmed from the quasi-experimental component. This section of this chapter contains the research questions and the statistical analysis they precipitated. Thus, the analysis for the achievement gap will precede that for the quasi-experimental component.

The achievement gap analysis relied on data that I downloaded from College Board for the years 2002, 2007, 2012, and 2017. Next, I analyzed the data for each of these years for each of the first three research questions. Thus, the analysis for the first three research questions will contain results for each of the four years.

Research question 1: what is the difference in the Calculus AB exam performance of Michigan students as a function of gender?

The results of the study indicated that the difference in the Calculus AB exam performance of Michigan students as a function of gender was statistically significant in each of the four years. To test the independence of Calculus AB exam scores from gender, I computed a 2 (gender) x 5 (exam score) Chi-square analysis using the Calculus AB exam scores for each of the four years. To provide additional context, I calculated descriptive statistics for each gender. Since the data is categorical, I calculated median and modal scores annually for each gender and tabulated frequency distributions for scores by gender for each year. For each of the four years that in the study, the annual median scores for each gender was 3. Table 7 contains the modes for each genders, and there was some variance in the modes. Table 8 through Table 11 contain the relative frequency distributions for each year, while the paragraphs below contain the statistical testing of the research hypothesis. The next paragraph contains a description of the modal scores, while the subsequent paragraph contains the results of the Chi-square analysis for each of the four years in the study.

In 2002, the mode was the same for both male and female participants. However, there was a stark contrast in the modal scores for both 2007 and 2012, as the mode for male participants was the highest score (5), while the mode for female participants was the lowest score (1). For 2017, the modal score for male participants remained at 5, while the modal score for females improved to 3. In summation, the modal score for males was higher than that of the females in three of the four years, and equal only in 2002, the earliest year of the study.

Table 7

Modal Scores by Gender

Year	Female	Male
2002	3	3
2007	1	5
2012	1	5
2017	3	5

Table 8

Relative Frequency of Scores by Gender for 2002

Exam Score	Female	Male
5	.1463	.2115
4	.2334	.2727
3	.2762	.3043
2	.1734	.1491
1	.1427	.0905

Table 8 contains the relative frequency distribution for Calculus AB exam scores for the year 2002. In 2002, male participants earned a higher percentage of each of the three highest scores, while female participants earned a higher percentage of each of the two lowest scores. To determine whether the difference was statistically significant, a Chi-square test was performed and it yielded χ^2 (4) =71.7151. The p-value was less than 0.00001; consequently, the result was statistically significant. The effect size for 2002 was small, as Cramer's V yielded a result of .12 (Bosco et al. 2015; Gignac & Szodorai, 2016).

Table 9

Relative Frequency of Scores by Gender for 2007

Exam Score	Female	Male
5	.1714	.2481
4	.1965	.2097
3	.2143	.1984
2	.1652	.1581
1	.2526	.1857

Presented in Table 9 is the relative frequency distribution for Calculus AB exam scores for 2007. In 2007, male participants earned a higher percentage of each of the two highest scores, while female participants earned a higher percentage of each of the three lowest scores. To determine whether the difference was statistically significant, a Chi-square test was performed and it yielded χ^2 (4) = 87.1576. The p-value was less than 0.00001; consequently, the result was statistically significant. The effect size for 2007 was small, as Cramer's V yielded a result of .11 (Bosco et al. 2015; Gignac & Szodorai, 2016).

Presented in Table 10 is the relative frequency distribution for Calculus AB exam scores for 2012. In 2012, male participants earned a higher percentage of each of the three highest scores, while female participants earned a higher percentage of each of the two lowest scores. To determine whether the difference was statistically significant, a Chi-square test was performed and it yielded χ^2 (4) = 93.1645. The p-value was less than 0.00001; consequently, the result was statistically significant. The effect size for 2012 was small, as Cramer's V yielded a result of .11 (Bosco et al. 2015; Gignac & Szodorai, 2016).

Table 10

Relative Frequency of Scores by Gender for 2012

Exam Score	Female	Male
5	.2014	.2828
4	.1875	.1893
3	.1886	.1939
2	.1158	.0879
1	.3066	.2461

Table 11

Relative Frequency of Scores by Gender for 2017

Exam Score	Female	Male
5	.1576	.2198
4	.1882	.2041
3	.2374	.2182
2	.2369	.2130
1	.1798	.1450

Table 11 contains the relative frequency distribution for Calculus AB exam scores for 2017. In 2017, male participants earned a higher percentage of each of the two highest scores, while female participants earned a higher percentage of each of the three lowest scores. To determine whether the difference was statistically significant, a Chi-square test was performed and it yielded $\chi^2(4) = 69.9074$. The p-value was less than 0.00001; consequently, the result was statistically significant. The effect size for 2017 was trivial, as Cramer's V yielded a result of .09 (Bosco et al. 2015; Gignac & Szodorai, 2016).

For research question 1, the researcher rejected the null hypothesis, as a statistically significant difference existed between genders in Calculus AB exam scores in each of the four years that were part of the study. Thus, males outscored females in each of the four years of the study. However, the effect size ranged from trivial to small, with monotonically decreasing effect sizes.

Research question 2: what is the difference in the Calculus AB exam performance of Michigan students as a function of ethnicity?

The results of the study indicated that the difference in the Calculus AB exam performance of Michigan students as a function of ethnicity was statistically significant in each of the four years. To test whether Calculus AB exam scores are independent from ethnicity, I computed 4 (ethnicity) x 5 (exam score) Chi-square analysis using the Calculus AB exam scores. To provide additional context, I calculated descriptive statistics for each of the four ethnicities, that is Asian, Black, Hispanic, and White. Because the data is categorical, I calculated median and annual modal scores for each ethnicity and tabulated relative frequency distributions for scores by ethnicity for each year. Table 12 contains the annual median score for each of the four ethnicities examined in this research question, while Table 13 contains the annual modal score for each ethnicity. The relative frequency distributions for the four year of the ethnic achievement analysis are in Table 14 through Table 17 below, while the accompanying paragraphs below contain the statistical testing of the research hypothesis for each year. The next two paragraphs contain descriptive statistics, after which the subsequent paragraphs contain the results of Chi-square analysis for each of the four years in the study.

Table 12

Ethnic Median Calculus AB score

Year	Asian	Black	Hispanic	White
2002	4	2	3	3
2007	4	1	2	3
2012	4	1	2	3
2017	4	2	2	3

As shown in Table 12, Asian participants had a higher median score in each of the four years, while White participants had the second highest median score for each of the four years. Hispanic participants tied for the second highest median in 2002, but had the third highest median each of the other years. Black participants had the lowest median in 2002, 2007, and 2012, but tied for lowest median in 2017.

Table 13

Ethnic Modal Calculus AB score

Year	Asian	Black	Hispanic	White
2002	4	2	3	3
2007	5	1	1	5
2012	5	1	1	1
2017	5	1	1	2

Similarly, Asian participants had the highest mode in each of the four years of the study, as shown in Table 13. White participants tied Asian participants for the highest mode in 2007, ranked second in both 2002 and 2017, and tied for the lowest mode in 2012. Hispanic participants tied for second highest mode in 2002, but tied for the lowest mode in 2007, 2012, and 2017. Black students had the lowest mode in all four years of the study.

Table 14

Relative Frequency of Scores by Ethnicity for 2002

Score	Asian	Black	Hispanic	White
5	.2812	.0448	.1512	.1809
4	.2958	.1791	.1628	.2571
3	.2347	.2612	.3372	.2912
2	.1198	.3060	.1512	.1574
1	.0685	.2090	.1977	.1134

Table 14 contains the relative frequency distribution for Calculus AB exam scores for 2002. In 2002, Asian participants earned the highest percentage of each of the top two scores, while Black participants earned the highest percentage of the lowest score. Asian participants earned the highest score at a ratio of more than 6:1 when compared with Black participants and at almost a rate of 2:1 when compared to Hispanic participants. White participants ranked second in earning the each of the two highest scores, while Hispanic participants had the second highest incidence of the lowest score. To determine whether the differences in ethnic performance were statistically significant, a Chi-square test was performed and it yielded $\chi^2(12) = 95.0022$. The p-value was less than 0.00001; consequently, the result was statistically significant. The effect size for 2002 was trivial, as Cramer's V yielded a result of .08 (Bosco et al. 2015; Gignac & Szodorai, 2016).

Table 15

Relative Frequency of Scores by Ethnicity for 2007

Score	Asian	Black	Hispanic	White
5	.2847	.0463	.1691	.2123
4	.2401	.0849	.1618	.2087
3	.2041	.1351	.1250	.2121
2	.1132	.1158	.1765	.1673
1	.1578	.6178	.3676	.1995

Table 15 contains the relative frequency distribution for Calculus AB exam scores for 2007. In 2007, Asian participants ranked first, White participants ranked second, and Hispanic participants ranked third in earning the highest percentage of each of the two highest scores. In contrast, Black participants were the least likely to earn each of the two highest scores and were the most likely to receive the lowest score. To clarify, Asian participants were almost twice as likely to earn a 5 as a 1, White participants were slightly more likely to earn a 5 than a 1, and Hispanic participants were about half as likely to earn a 5 than 1. However, Black participants were fifteen times more likely to earn a 1 than a 5; and, from 2002 to 2007, the rate for Blacks of scoring a 1 more than tripled from 20.9% to 61.8%. To determine whether the differences in ethnic performance were statistically significant, a Chi-square test was performed and it yielded $\chi^2(12) = 321.6857$. The p-value was less than 0.00001; consequently, the result was statistically significant. The effect size for 2007 was small, as Cramer's V yielded a result of .13 (Bosco et al. 2015; Gignac & Szodorai, 2016).

Table 16

Relative Frequency of Scores by Ethnicity for 2012

Score	Asian	Black	Hispanic	White
5	.3991	.0643	.1190	.2393
4	.1861	.0924	.1714	.1957
3	.1498	.0643	.1714	.2034
2	.0946	.0803	.0524	.1031
1	.1703	.6988	.4857	.2584

Table 16 contains the relative frequency distribution for Calculus AB exam scores for 2012. In 2012, Asian participants ranked first, White participants ranked second, and Hispanic participants ranked third in earning the highest percentage of the highest score. In contrast, Black participants were the least likely to earn the highest score and were the most likely to receive the lowest score. To clarify, Asian participants were more than twice as likely to earn a 5 as a 1 and White participants were almost as likely to earn a 5 as a 1. However, Hispanic participants were four times more likely to score a 1 than a 5 and Black participants were about eleven times more likely to score a 1 than a 5. From 2007 to 2012, the rate of Black participants scoring a 1 increased by 8 percentage points, or 13%, while the rate of Hispanic participants scoring a 1 increased by about 12 percentage points, or 32%. To determine whether the differences in ethnic performance were statistically significant, a Chi-square test was performed and it yielded χ^2 (12) = 391.4905. The p-value was less than 0.00001; consequently, the result was statistically significant. The effect size for 2012 was small, as Cramer's V yielded a result of .13 (Bosco et al. 2015; Gignac & Szodorai, 2016).

Table 17

Relative Frequency of Scores by Ethnicity for 2017

Score	Asian	Black	Hispanic	White
5	.2984	.0489	.1282	.1857
4	.2105	.0703	.1256	.2081
3	.2261	.1621	.2256	.2301
2	.1615	.2355	.2564	.2309
1	.1036	.4832	.2641	.1453

Table 17 contains the relative frequency distribution for Calculus AB exam scores for 2017. In 2017, Asian participants ranked first, White participants ranked second, and Hispanic participants ranked third in earning the highest percentage of each of the two highest scores. In contrast, Black participants were the least likely to earn the highest score and were the most likely to receive the lowest score. To clarify, Asian participants were almost three times more likely to earn a 5 than a 1 and White participants were more likely to earn a 5 than a 1. However, Hispanic participants were more than twice as likely to score a 1 as a 5 and Black participants were about ten times more likely to score a 1 than a 5. Additionally, White participants were more than 3 times more likely to score a 5 than Black participants were and one and a half times more likely than Hispanic participants to score a 5. To determine whether the differences in ethnic performance were statistically significant, a Chi-square test was performed and it yielded $\chi^2(12) = 430.8514$. The p-value was less than 0.00001; consequently, the result was statistically significant. The effect size for 2017 was small, as Cramer's V yielded a result of .13 (Bosco et al. 2015; Gignac & Szodorai, 2016).

For research question 2, the researcher rejected the null hypothesis, as a statistically significant difference existed between ethnicities in Calculus AB scores in each of the four years that was part of the study. Based on Cohen's criteria, the effect size for the ethnic achievement gap analysis was trivial in 2002. However, the effect size increased from .08 in 2002 to .13 for each of the years 2007, 2012, and 2017, which translates to a small effect size (Bosco et al. 2015; Gignac & Szodorai, 2016).

Research question 3: what is the difference in the Calculus AB exam performance of Michigan's Hispanic students as a function of gender?

The results of the study indicated that the difference in the Calculus AB exam performance of Michigan's Hispanic students as a function of gender was statistically significant in only the most recent of the four years that were tested, that is 2017. To test whether Calculus AB exam scores are independent from gender for Hispanic participants, I computed 2 (ethnicity) x 5 (exam score) Chi-square analysis using the Calculus AB exam scores for each of the four years. To provide additional context, I calculated descriptive statistics for each gender. Since the data is categorical, I calculated median and modal scores annually for each gender and tabulated relative frequency distributions for scores by gender for each year. Since the data is categorical, I calculated the annual median and modal scores for each gender and I tabulated the relative frequency distributions for scores by ethnicity for each year. The annual median scores were the same for each gender; the median score for each gender was 3 in 2002 and 2 in each of the other three years. Table 18 contains the modes for each gender for Hispanic students, and there was some variance in the modes. The relative frequency distributions for the four year of the Hispanic gender analysis are in Table 19 through Table 22 below, while the accompanying paragraphs below contain the statistical testing of the research hypothesis for each year. Contained in the next two paragraphs are the descriptive statistics, after which the subsequent paragraphs contain the results of Chi-square analysis for each of the four years in the study.

Table 18

Hispanic Modal Scores by Gender

Year	Female	Male
2002	1	3
2007	1	1
2012	1	1
2017	1/2	3

Table 18 contains the modal scores for Hispanic participants by gender for 2002. In 2002, the mode for Hispanic male participants was 3, while the mode for Hispanic female participants was 1. In both 2007 and 2012, the mode was 1 for both Hispanic male participants and Hispanic female participants. However, in 2017, the modal score for Hispanic male participants increased to 3, while the modal score for Hispanic females was both 1 and 2. In summation, the modal score for Hispanic males was higher than that of Hispanic females in two of the four years, while it was the same for both Hispanic genders in the other two years.

Table 19 contains the relative frequency distribution for Calculus AB exam scores for Hispanic participants for 2002. In 2002, Hispanic female participants earned a higher percentage of each of both the highest and the lowest scores, while Hispanic male participants earned a higher percentage of the middle three scores. While Hispanic female participants were more likely to score a 5 than Hispanic male participants by a ratio of 3:2, Hispanic female participants were more likely to score a 1 than Hispanic male participants by a ratio of 5:2. Therefore, while Hispanic male participants were equally likely to score a 1 or a 5, Hispanic female participants were more likely to score a 1 than a 5 by a ratio of approximately 3:2. To determine whether the difference was statistically significant, a Chi-square test was performed and it yielded $\chi^2(4) = 6.3785$. The p-value was .17261; consequently, the result was not statistically significant. The effect size for 2002 was small, as Cramer's V yielded a result of .27 (Bosco et al. 2015; Gignac & Szodorai, 2016).

Table 19

Relative Frequency of Scores by Gender for Hispanics in 2002

Exam Score	Female	Male
5	.1892	.1224
4	.1622	.1633
3	.2432	.4082
2	.1081	.1837
1	.2973	.1224

Table 20 contains the relative frequency distribution for Calculus AB exam scores for Hispanic participants for 2007. In 2007, Hispanic male participants earned a higher percentage of the highest score, while Hispanic female participants earned a higher percentage of the lowest scores. While Hispanic females were nearly three times more likely to score a 1 than a 5, Hispanic male participants were less than twice as likely to score a 1 as a 5. To determine whether the difference was statistically significant, a Chi-square test was performed and it yielded $\chi^2(4) = 3.0196$. The p-value was .554555; consequently, the result was not statistically significant. The effect size for 2007 was small, as Cramer's V yielded a result of .15 (Bosco et al. 2015; Gignac & Szodorai, 2016).

Table 20

Relative Frequency of Scores by Gender for Hispanics in 2007

Exam Score	Female	Male
5	.1481	.1928
4	.1667	.1566
3	.0926	.1446
2	.1481	.1928
1	.4444	.3133

Table 21

Relative Frequency of Scores by Gender for Hispanics in 2012

Exam Score	Female	Male
5	.1515	.0901
4	.1414	.1982
3	.1818	.1622
2	.0505	.0541
1	.4747	.4955

Table 21 contains the relative frequency distribution for Calculus AB exam scores for Hispanic participants for 2012. In 2012, Hispanic female participants earned a higher percentage of the highest score, while Hispanic male participants earned a higher percentage of the lowest scores. While Hispanic females were nearly three times more likely to score a 1 than a 5, Hispanic male participants were in excess of five times more likely to score a 1 than a 5. To determine whether the difference was statistically significant, a Chi-square test was performed and it yielded $\chi^2(4) = 2.8196$. The p-value was .588449; consequently, the result was not statistically significant. The effect size for 2012 was small, as Cramer's V yielded a result of .12 (Bosco et al. 2015; Gignac & Szodorai, 2016).

Table 22

Relative Frequency of Scores by Gender for Hispanics in 2017

Exam Score	Female	Male
5	.0769	.1731
4	.1264	.1250
3	.1923	.2548
2	.3022	.2163
1	.3022	.2308

Table 22 contains the relative frequency distribution for Calculus AB exam scores for Hispanic participants for 2017. In 2017, Hispanic male participants earned a higher percentage of the highest score, while Hispanic female participants earned a higher percentage of the lowest scores. While Hispanic females were nearly four times as likely to score a 1 as a 5, Hispanic male participants were only four thirds as likely to score a 1 as a 5. To determine whether the difference was statistically significant, a Chi-square test was performed and it yielded $\chi^2(4) = 13.3472$. The p-value was .009698; consequently, the result was statistically significant. The effect size for 2017 was small, as Cramer's V yielded a result of .19 (Bosco et al. 2015; Gignac & Szodorai, 2016).

For research question 3, the researcher rejected the null hypothesis for the first three years, as a statistically significant difference did not exist between genders in Calculus AB scores in 2002, 2007, and 2012. In contrast, the researcher rejected the null hypothesis for 2017, as a statistically significant difference existed between genders in Calculus AB scores for that year. For each year of the Hispanic gender gap analysis, the effect size was small.

Research question 4: what is the difference in Calculus AB exam performance of students as a function of a yearlong two-hour AP Calculus class versus a one-hour AP Calculus class?

The results of the study indicated that the difference in Calculus AB exam performance of students as a function of a yearlong two-hour AP Calculus class versus a one-hour AP Calculus class was statistically significant. The design of this study included analysis of data for research questions four through seven with Chi-square analysis, but the expected value of too many of the subgroups was below five. This made Chi-square analysis unreliable, so analysis for research questions four through seven a Fisher-Freeman-Halton exact probability test replaced Chi-square analysis. However, this change did not require any additional changes to the methodology.

Thus, effect size was still calculated using Cramer's V and subsequently evaluated using Cohen's criteria.

To determine whether the comparison and experiential groups were statistically similar, I compared the SAT math scores of the comparison and experiential groups. For participants who took the ACT instead of the SAT, I converted ACT math scores to SAT math scores using official College Board concordance tables. Then, scores were grouped into hundred point categories and the resultant distribution of scores was compared with a 2 (groups) x 5 (SAT categories) Fisher-Freeman-Halton exact probability test. The Fisher-Freeman-Halton test yielded a p-value of 0.137; consequently, the SAT category distributions for the comparison group and the experiential group were not statistically significantly different. The effect size was small, as Cramer's V yielded a result of .21 (Bosco et al. 2015; Gignac & Szodorai, 2016).

The comparison group received one hour of Calculus instruction in a yearlong course and the experiential group received two hours of daily Calculus instruction in a yearlong course. Following a year of instruction, all students in both groups took the Calculus AB. Next, a 2 (treatments) x 5 (exam score) Fisher-Freeman-Halton test was computed using the Calculus AB exam scores to determine whether Calculus AB exam scores were independent from the length of the instructional period. The Fisher-Freeman-Halton test yielded a p-value less than .001; consequently, the result was statistically significant. The effect size was medium, as Cramer's V yielded a result of .48 (Bosco et al. 2015; Gignac & Szodorai, 2016).

For research question 4, the researcher rejected the null hypothesis, as a statistically significant difference existed in the Calculus AB exam scores between participants in the comparison and the experiential groups. The effect size for the difference between the experiential treatment and the comparison treatment was medium. These results suggest that the experiential treatment of two hours of Calculus instruction produced higher Calculus AB exam achievement than the comparison treatment of a single hour of daily Calculus instruction for all students.

Research question 5: what is the difference in Calculus AB exam performance of Hispanic students as a function of a yearlong two-hour AP Calculus class versus a one-hour AP Calculus class?

The results of the study indicated that the difference in Calculus AB exam performance of Hispanic students as a function of a yearlong two-hour AP Calculus class versus a one-hour AP Calculus class was statistically significant. To determine whether the comparison and experiential groups were statistically similar, I compared SAT math scores of the comparison and experiential groups for just the Hispanic participants. For participants who took the ACT instead of the SAT, I converted ACT math scores to SAT math scores using official College Board concordance tables. Then, scores were grouped into hundred point categories and the resultant distribution of scores was compared with a 2 (groups) x 5 (SAT categories) Fisher-Freeman-Halton exact probability test. The Fisher-Freeman-Halton test yielded a p-value of 0.272; consequently, the SAT category distributions for the comparison group and the experiential group were not statistically significantly different. The effect size was small, as Cramer's V yielded a result of .21 (Bosco et al. 2015; Gignac & Szodorai, 2016).

The comparison group received one hour of Calculus instruction in a yearlong course and the experiential group received two hours of daily Calculus instruction in a yearlong course. Following a year of instruction, all students in both groups took the Calculus AB. A 2 (treatments) x 5 (exam score) Fisher-Freeman-Halton test was computed using the Calculus AB exam scores of only the Hispanic students to determine whether Calculus AB exam scores were independent from the two treatments. The Fisher-Freeman-Halton test yielded a p-value that was less than .001; consequently, the result was statistically significant. The effect size was medium, as Cramer's V yielded a result of .49 (Bosco et al. 2015; Gignac & Szodorai, 2016).

For research question 5, the researcher rejected the null hypothesis, as a statistically significant difference existed in Calculus AB exam scores of Hispanic students between participants in the comparison and the experiential groups. The effect size for the difference between the experiential treatment and the comparison treatment was medium. These results suggest that the experiential treatment of two hours of Calculus instruction produced higher Calculus AB exam achievement than the comparison treatment of a single hour of daily Calculus instruction for Hispanic students.

Research question 6: what is the difference in Calculus AB exam performance of Hispanic male students as a function of a yearlong two-hour AP Calculus class versus a one-hour AP Calculus class?

The results of the study indicated the difference in Calculus AB exam performance of Hispanic male students as a function of a yearlong two-hour AP Calculus class versus a one-hour AP Calculus class was statistically significant. To determine whether the comparison and experiential groups were statistically similar, I compared SAT math scores of the comparison and experiential groups for just the Hispanic male participants. For participants who took the ACT instead of the SAT, I converted ACT math scores to SAT math scores using official College Board concordance tables. Then, scores were grouped into hundred point categories and the resultant distribution of scores was compared with a 2 (groups) x 5 (SAT categories) Fisher-Freeman-Halton exact probability test. The Fisher-Freeman-Halton test yielded a p-value of 0.707; consequently, the SAT category distributions for the comparison group and the experiential group were not statistically significantly different. The effect size was small, as Cramer's V yielded a result of .19 (Bosco et al. 2015; Gignac & Szodorai, 2016).

The comparison group received one hour of Calculus instruction in a yearlong course and the experiential group received two hours of daily Calculus instruction in a yearlong course. Following a year of instruction, all students in both groups took the Calculus AB. A 2 (treatments) x 5 (exam score) Fisher-Freeman-Halton test was computed using the Calculus AB exam scores of only Hispanic students who indicated their gender as male to determine whether Calculus AB exam scores were independent from the two treatments. The Fisher-Freeman-Halton test yielded a p-value that was less than .001; consequently, the result was statistically significant. The effect size was large, as Cramer's V yielded a result of .63 (Bosco et al. 2015; Gignac & Szodorai, 2016).

For research question 6, the researcher rejected the null hypothesis, as a statistically significant difference existed in Calculus AB exam scores of Hispanic male students between participants in the comparison and the experiential groups. The effect size for the difference between the experiential treatment and the comparison treatment was large. These results suggest that the experiential treatment of two hours of Calculus instruction produced higher Calculus AB exam achievement than the comparison treatment of a single hour of daily Calculus instruction for Hispanic males.

Research question 7: what is the difference in Calculus AB exam performance of Hispanic female students as a function of a yearlong two-hour AP Calculus class versus a one-hour AP Calculus class?

The results of the study indicated the difference in Calculus AB exam performance of Hispanic female students as a function of a yearlong two-hour AP Calculus class versus a one-hour AP Calculus class was statistically significant. To determine whether the comparison and experiential groups were statistically similar, I compared SAT math scores of the comparison and experiential groups for just the Hispanic female participants. For participants who took the ACT instead of the SAT, I converted ACT math scores to SAT math scores using official College Board concordance tables. Then, scores were grouped into hundred point categories and the resultant distribution of scores was compared with a 2 (groups) x 5 (SAT categories) Fisher-Freeman-Halton exact probability test. The Fisher-Freeman-Halton test yielded a p-value of 0.099; consequently, the SAT category distributions for the comparison group and the experiential group were not statistically significantly different. The effect size was medium, as Cramer's V yielded a result of .31 (Bosco et al. 2015; Gignac & Szodorai, 2016).

The comparison group received one hour of Calculus instruction in a yearlong course and the experiential group received two hours of daily Calculus instruction in a yearlong course. Following a year of instruction, all students in both groups took the Calculus AB. A 2 (treatments) x 5 (exam score) Fisher-Freeman-Halton test was computed using the Calculus AB exam scores of only Hispanic students who indicated their gender as female to determine whether Calculus AB exam scores were independent from the two treatments. The Fisher-Freeman-Halton test yielded a p-value that was less than .001; consequently, the result was statistically significant. The effect size was medium, as Cramer's V yielded a result of .45 (Bosco et al. 2015; Gignac & Szodorai, 2016).

For research question 7, the researcher rejected the null hypothesis, as a statistically significant difference existed in Calculus AB exam scores of Hispanic female students between participants in the comparison and the experiential groups. The effect size for the difference between the experiential treatment and the comparison treatment was medium. These results suggest that the experiential treatment of two hours of Calculus instruction produced higher Calculus AB exam achievement than the comparison treatment of a single hour of daily Calculus instruction for Hispanic females.

Summary

The purpose of this study was to determine whether there were gender and ethnic achievement gaps in the state of Michigan and to test whether increasing the length of the instructional period from one hour to two hours might remedy the problem. To that end, there were two components to this study. The first component of the study was an achievement gap analysis, while the second component was a quasi-experimental component that tested whether the hypothesis of increasing the instructional period might result in an increase in Calculus AB exam scores, while providing a potential remedy to the problem of achievement gaps.

The achievement gap analysis consisted of three parts. The first part was the gender gap analysis. This part of the study compared the Calculus AB exam scores of males and females in the years 2002, 2007, 2012, and 2017. Regarding the gender achievement gap analysis, the results of the analysis indicated that a statistically significant difference existed for all four years with males outscoring females each year. The second part of the achievement gap analysis was an ethnic gap analysis. This part of the study consisted of a comparison of Calculus AB exam scores for students of Asian, Black, Hispanic, and White ethnicities for the years 2002, 2007, 2012, and 2017. For the ethnic achievement gap analysis, a statistically significant difference was found in the distribution of scores for the four ethnicities in all four years that were analyzed, with Asian and White students outscoring Hispanic and Black students in each year. Regarding the Hispanic gender achievement gap analysis, the results of the analysis indicated that there was a statistically significant difference in the distribution of scores only in the most recent year of the study (2017). Therefore, the researcher rejected the null hypotheses for all four years in both the gender and ethnic achievement gap analyses, whereas, for the Hispanic gender achievement gap analysis, the researcher rejected the null hypothesis only for 2017.

The quasi-experimental component produced four research questions. In an overall comparison of the scores of the comparison and the experiential groups, a statistically significant difference existed in the Calculus AB exam scores of the participants of the two groups. In a comparison of the Calculus AB exam scores of only the Hispanic students, a statistically significant difference existed between participants in the comparison and the experiential groups. In a comparison of the Calculus AB exam scores of only the Hispanic male participants, a statistically significant difference existed between participants in the comparison and the experiential groups. Similarly, in a comparison of the Calculus AB exam scores of the just the Hispanic female participants, a statistically significant difference existed between the participants in the comparison and the experiential group. Effect sizes for the quasi-experimental component ranged from medium to large. Consequently, the researcher rejected all four of the null hypotheses in the quasi-experimental component.

Chapter 5: Discussion and Conclusion

The purpose of this study was to investigate whether there were achievement gaps on the Calculus AB exam in Michigan and to determine whether increasing the length of the daily instructional period might offer an opportunity for schools to address those achievement gaps. The three achievement gaps on the Calculus AB exam that were investigated in the state of Michigan were gender, ethnic, and gender within the Hispanic population. Following the investigation of Michigan's achievement gaps on the Calculus AB exam, I conducted a quasi-experiment to determine whether increasing the length of daily instruction from one hour to two might result in a statistically significant difference in the distribution of Calculus AB exam scores. This chapter contains an analysis of the results from both components of the study and discussion of those results in light of the literature. Additionally, this chapter contains suggestions for schools, districts, and educational policy makers. Consequently, the organization of this chapter includes eight sections, beginning with this (a) introduction. The subsequent sections of this chapter are entitled (b) summary of the results, (c) discussion of the results, (d) discussion of the results in relation to the literature, (e) limitations, (f) implications of the results for practice, policy, and theory, (g) recommendations for further research, and (h) conclusion.

Summary of the Results

This study contained seven research questions connected to the two components. The first three research questions stemmed from the achievement gap analysis, while the other four research questions emerged from the quasi-experiment. Below are the seven research questions.

1. What is the difference in the Calculus AB exam performance of Michigan students as a function of gender?

2. What is the difference in the Calculus AB exam performance of Michigan students as a function of ethnicity?

3. What is the difference in the Calculus AB exam performance of Michigan's Hispanic students as a function of gender?

4. What is the difference in Calculus AB exam performance of students as a function of a yearlong two-hour AP Calculus class?

5. What is the difference in Calculus AB exam performance of Hispanic students as a function of a yearlong two-hour AP Calculus class?

6. What is the difference in Calculus AB exam performance of Hispanic male students as a function of a yearlong two-hour AP Calculus class?

7. What is the difference in Calculus AB exam performance of Hispanic female students as a function of a yearlong two-hour AP Calculus class?

This study, grounded in OTL theory, was an investigation into whether the principles of OTL might offer a possible means of addressing gender achievement gaps and ethnic achievement gaps on the Calculus AB exam. While the exact elements that contribute to OTL are still under debate, seemingly every model includes the indispensable variable of instructional time and content (Elliott & Bartlett, 2016). Consequently, the focal point for this investigation was the variable of time because there is evidence that yearlong instruction in block schedule Calculus classes may increase achievement (Gullatt, 2006; Keen, 1996; Smith & Camara, 1998).

Researchers have concluded that there is a nationally ubiquitous Advanced Placement gender achievement gap (Melsom, 2008; Moore & Slate, 2010, 2011; Moore et al., 2012; Morris, 2013; Morris et al., 2015). Likewise, the United States has an ethnic achievement gap when it comes to Advanced Placement exam scores (Holmes, 2013; Jara, 2013). Researchers, in previous studies, have called for future research to conduct single state analyses of Advanced Placement gender achievement differences and gender differences within ethnic groups (Jara, 2013; Morris, 2013). Therefore, the first component of the study was an achievement gap analysis for the state of Michigan, which included a gender achievement gap analysis, an ethnic achievement gap analysis, and a Hispanic gender achievement gap analysis. The purpose of this component was to conduct a current analysis of Advanced Placement achievement gaps for a locale that had not yet been analyzed, which additionally establishes the necessity for investigation into strategies that may lead to equity for all students.

Regarding the gender achievement gap analysis, a statistically significant difference existed between the Calculus AB exam scores of the two genders in each of the four years that were tested. Regarding the ethnic gap analysis, a statistically significant difference existed between the Calculus AB exam scores of the four ethnicities in each of the four years analyzed in the course of this study. Regarding the Hispanic gender achievement gap analysis, a statistically significant difference in Calculus AB exam scores for the two genders only existed in 2017. Consequently, within the state of Michigan, there exists both a gender achievement and an ethnic achievement gap; and, within the last five years, a Hispanic gender achievement gap has developed.

The second component of the study was a quasi-experiment. The comparison group were participants who took the Advanced Placement exam in either 2014 or 2015. The experiential group consisted of participants who took the Advanced Placement exam in either 2016 or 2017. The comparison group received one hour of Calculus instruction daily, whereas the experiential group received two hours of instruction daily. The quasi-experiment included four comparisons of the comparison group and the experiential group. The first comparison was for the overall

Calculus AB exam results. The second comparison was for the Calculus AB exam results of only the Hispanic students. The third comparison was for the Calculus AB exam results of only Hispanic males. The fourth comparison was for the Calculus AB exam results for only Hispanic females. Regarding the results from the quasi-experiment, a statistically significant difference existed between participants in the comparison group and the experiential group in all four comparisons.

Many researchers have conducted studies on gender achievement differences pertaining to the Advanced Placement exam. Similarly, there are many studies in the literature whose subject is ethnic achievement differences on Advanced Placement exams. So although researchers have conducted many studies to investigate either gender differences or ethnic differences in Advanced Placement exam performance, there have been very few studies whose subject was solving the problem of achievement gaps on those exams. The results of this study confirm that both the gender achievement gaps and ethnic achievement gaps that exist in various locales are also present in Michigan. More significantly, this study fills a void in the literature as it investigated a potential solution to the problem of Advanced Placement achievement gaps.

Discussion of the Results

This section of this chapter contains discussion of the results from Chapter 4. This section contains four subsections. The first subsection contains discussion pertaining to the gender achievement gap analysis. The second subsection contains discussion pertaining to the ethnic achievement gap analysis. The third subsection contains discussion pertaining to the Hispanic gender achievement gap analysis. The fourth section contains discussion pertaining to the quasi-experiment.

Gender achievement gap. Evidence from this study suggests that a persistent gender achievement gap on the Calculus AB exam has existed in the state of Michigan over the past 15 years, with boys consistently outscoring girls. During the course of the study, the gender distribution in population in Michigan's schools has remained relatively unchanged, as 51.4% of all students in Michigan were male in 2002 and 51.5% of all students in Michigan were male in 2017 (College Board, 2002, 2017c; Michigan Department of Education, 2018). Consequently, the male dominance in the Advanced Placement participation rate in Michigan, which has declined from 54.3% male in 2002 to 53.2% male in 2017, seems to be approaching the state population distribution. While both more males and females are electing to sit for the Advanced Placement exam, as male participation has increased from 2,629 students in 2002 to 4,400 students in 2017, the increase in participation is not translating to elimination of the gender achievement gap (College Board, 2002, 2017c). On the Calculus AB exam, male students continue to post a higher mode, higher percentages of scoring 5, and lower percentages of scoring 1; all of which supports the conclusion that the achievement gap continues.

College Board (2017b) produces both the Advanced Placement exams and the SAT. There has been a decades-long gender achievement gap on the SAT (Perry, 2016). Cimpian, Lubienski, Timmer, Makowski, and Miller (2016) conclude that gender achievement gaps in mathematics emerge early in elementary school and plague a cohort from kindergarten to graduation. Prior knowledge is crucial to mathematics learning, as foundational skills mount throughout the journey toward graduation. As the capstone of high school mathematics, Calculus requires considerable prior knowledge. If there is an annual discrepancy between the two genders in the amount of prior learning accumulated by students enrolling in Calculus, then the conclusion that a gender achievement gap exists in Calculus AB exams is hardly surprising.

Ethnic achievement gap. A persistent ethnic achievement gap has also existed on the Calculus AB exam over the past 15 years in Michigan. However, there is also a participation gap between the ethnicities on the Calculus AB exam. In 2002, White students accounted for 72.9% of all students in Michigan schools, but they account for 86.3% of students taking the Calculus AB exam; meanwhile, Asian students, who only accounted for 1.9% of all students in Michigan schools, made up 8.9% of all students taking the Calculus AB exam (College Board, 2002; Michigan Department of Education, 2018). In 2002, a parallel underrepresentation of Black students, who make up 19.7% of all students in Michigan, but only 2.9% of students taking the Calculus AB exam, accompanied an overrepresentation of Asian and White students (College Board, 2002; Michigan Department of Education, 2018). Hispanic students, who made up 3.8% of all students in Michigan schools, constituted only 1.9% of students sitting for the Calculus AB exam (College Board, 2002; Michigan Department of Education, 2018).

The participation gap still existed in 2017, although the percentages have changed. In 2017, White students were 66.6% of the overall school population in Michigan, but made up 79.9% of students taking the Calculus AB exam; while, Asian students accounted for 3.3% of the total school population, but made up 11.2% of students sitting for the Calculus AB exam (College Board, 2017c; Michigan Department of Education, 2018). The overrepresentation of Whites and Asians taking the Calculus AB exam was still accompanied by an underrepresentation of Black and Hispanic students, as Black students accounted for 18.0% of all students in Michigan, but made up 4.1% of students taking Calculus AB exam and Hispanic students accounted for 7.7% of all students in Michigan, but only made up 4.9% of students taking the Calculus AB exam (College Board, 2017c; Michigan Department of Education, 2018).

Although in 2017 there were more Black and Hispanic students taking the Calculus AB exam, the percentage of those ethnicities scoring a 1 increased dramatically between 2002 and 2017.

The ethnic achievement gap analysis showed that Calculus AB exam scores seem to be a function of ethnicity, as Asian students outscored White students, who outscored both Hispanic and Black students. That is, Asian students outscored White students on the Calculus AB exam for each of the years that were examined between 2002 and 2017; meanwhile, White students outscored both Black and Hispanic students during each of the years that were examined. In the course of this study, I did not attempt to uncover the reason for the discrepancy that existed between the ethnicities, but the literature indicated that it socioeconomic status and parental education levels are likely factors causing the gaps (Alford-Stephens, 2016; Champion & Mesa, 2016; Hernandez, 2014; Kodippili, 2011).

Hispanic gender achievement gap. The Hispanic gender analysis showed that male and female students who considered themselves Hispanic scored similarly on the Calculus AB exam in 2002, 2007, and 2012. In contrast, the 2017 analysis showed that a statistically significant discrepancy between Hispanic males and their female counterparts had developed in performance on the Calculus AB exam. The methodology used in this study did not provide a means to determine why there was no gender achievement gap for Hispanic students in 2002, 2007, and 2012; nor was there any attempt made to determine why a gender achievement gap emerged in 2017. However, the sample size was extremely small in 2002, and still relatively small in 2007 and 2012, perhaps the Hispanic students that were choosing to take the Calculus AB exam in those years were equally exceptional. The number of Hispanic test takers increased dramatically by 2017, which seems to have been enough for the Hispanic population to exhibit the gender discrepancy that the entire population already displayed. Additionally, the number of

Hispanic test takers remained small enough that the addition of an Advanced Placement Calculus program in one or two largely Hispanic communities could be enough to introduce gender differences into the Calculus AB exam results.

Results from the quasi-experimental component. Regarding the quasi-experimental component, statistically significant differences in Calculus AB exam score distribution existed in each of the four comparisons between the comparison group and the experiential group. While the number of participants in this study was only a fraction of the students in Michigan that take the Calculus AB exam annually, this study included approximately 10% of Michigan's Hispanic students who took the Calculus AB exam from 2013 to 2017. Interpreted using Cohen's criteria, the effect size of the differences between the experiential group and the comparison group were medium for the overall group of students, Hispanic students, and Hispanic female students; while, the effect size was large for Hispanic male students (Bosco et al. 2015; Gignac & Szodorai, 2016). With the difference in effect size between Hispanic male students and Hispanic female students, together with the small population size of Hispanic test takers in Michigan, it is possible that the quasi-experimental component of this study contributed to the emergence of a statistically significant difference in the distribution of Calculus AB exam scores in the Hispanic gender achievement gap analysis. Although there was a statistically significant difference in Calculus AB exam performance between the comparison group and the experiential groups as a whole, there were arguably too few non-Hispanic students within this study to conclude that this particular result was conclusive. However, the results of the quasi-experimental component of the study seem to indicate that increasing the amount of instructional time for Advanced Placement Calculus classes may increase Calculus AB exam scores for Hispanic students, Hispanic male students, and Hispanic female students.

Discussion of the Results in Relation to the Literature

The results of the gender achievement gap analysis indicated that a statistically significant gender achievement gap exists in the state of Michigan pertaining to the Calculus AB exam, as male students outscored female participants in every year of the analysis. Researchers have concluded that there exists a gender achievement gap on the Calculus AB exam in favor of male students (Melsom, 2008; Moore & Slate, 2010, 2011; Morris & Slate, 2012; Morris, 2013; Morris et al., 2015). Evidence exists that gender achievement gaps develop in mathematics early on and continue to graduation (Cimpian et al., 2016). Advanced Placement Calculus is usually the last high school mathematics course for those who choose to take it, so the existence of a gender achievement gap on the Calculus AB exam may be a result of the accretion of achievement differences throughout the course of elementary, middle, and high school. Thus, the results of the gender gap analysis were consistent with the literature (Melsom, 2008; Moore & Slate, 2010, 2011; Morris & Slate, 2012; Morris, 2013; Morris et al., 2015).

The results of the ethnic achievement gap analysis indicated that a statistically significant ethnic achievement gap exists in the state of Michigan pertaining to the Calculus AB exam in all four years analyzed in the course of this study. Researchers have concluded that there exists an ethnic achievement gap on the Calculus AB exam in favor of Asian and White students at the expense of Black and Hispanic students (Holmes, 2013; Jara, 2013). Flores (2007) posits that ethnic achievement gaps may be the result of discrepancies that exist in OTL. For example, Black and Hispanic students are less likely to received early educational opportunities and more likely to have inexperienced or underqualified teachers than their White and Asian peers (Corcoran & Evans, 2008; Darling-Hammond, 2010; Flores, 2007). Darling-Hammond (2010) ascribes the differences in OTL for minority students to the resegregation of the public school system, which has meant that some 40% of minority students attend predominantly minority schools. Resegregation has led to a lack of resources for many schools that serve a minority populace that is frequently impoverished (Darling-Hammond, 2010). Whether the differences in Calculus AB exam achievement in Michigan is a phenomenon that is a byproduct of discrepancies in OTL remains unclear. Regardless, the results of the ethnic achievement gap analysis were also consistent with the literature.

The results of the Hispanic gender achievement gap analysis indicated that a statistically significant Hispanic gender achievement gap existed in the state of Michigan pertaining to the Calculus AB exam in only one of the four years examined in this study, 2017. While gender studies tend to be ubiquitous, there is not much literature investigating the phenomenon of a Hispanic gender achievement gap on Advanced Placement exams in general or the Calculus AB exam in particular. In one of the few studies examining both gender and ethnicity whose subject was Advanced Placement achievement, Holmes (2013) documents a similar change in the national Advanced Placement scores for Hispanic students, noting a negative trend in Advanced Placement scores for Hispanic female students. In addition, Hispanic females outperformed Hispanic males in Advanced Placement achievement every year from 1997 to 2004 (Holmes, 2013). Nationally, Hispanic males outscored their female counterparts every year from 2005 to 2012 (Holmes, 2013). In the state of Michigan, males did not score statistically significantly different from females in 2012 on the Calculus AB exam, but the shift toward higher scores for males was evident by 2017. No attempt to explain the reason for the change in scoring by gender was present in this study; however, the results of this study reflect the national trends reported by Holmes (2013).

The investment of time is an important factor that affects student achievement (Elliott & Bartlett, 2016; Floden, 2002; Ottmar et al., 2013). Studies have shown that students who took two semesters of mathematics in a block setting performed better on the Calculus AB exam than those who took only a single semester block of Calculus, regardless of whether the first semester consisted of a prerequisite mathematics course or Calculus itself (Hansen et al., 2000; Lightner, 2009). The results of the quasi-experimental component were consistent with this finding, as all four comparisons between the experiential group and the comparison group fared statistically significantly better with two hours of yearlong Calculus instruction. Consistent with OTL theory, an additional hour of Calculus instruction daily resulted in higher Calculus AB exam achievement for the following groups of participants: Hispanic students, Hispanic males, Hispanic females, and all students. While the investment of extra time did improve performance, it remains unclear whether this strategy would help to close either the gender achievement gap or the ethnic achievement gap.

Limitations

One limitation of the study was that the researcher was also the instructor of the Advanced Placement Calculus class for each of the participants in the quasi-experimental component of the study. The involvement of the researcher in the study itself introduces the possibility of observer bias (Mahtani, Spencer, Brassey, & Heneghan, 2018). Since the researcher is interacting with the subjects, observer bias was virtually unavoidable in this study (Mahtani et al., 2018).

A second limitation of this study was due to a lack of a metric to quantify the impact of the non-temporal variables of OTL, as the study was restricted to a single dimension of OTL. The Carroll Model includes five dimensions in OTL, which means that I excluded four of the dimensions of OTL (Elliott & Bartlett, 2016; Witt et al., 2015). In example, the study included an assumption that the content of the course remained the same. Although I used the same course syllabus for each of the four years of the course, for the experiential group there were more minutes to fill in each day. Consequently, students in the experiential group were probably more examples and the likely performed more exercises in class, which may have resulted in an unintended difference in either the breadth or depth of content taught. Another assumption included in this study was that instructional quality remained constant over the four years; although it is reasonable to assume the instructor became more proficient at explaining various concepts.

A third limitation of this study was the inability to compute the exact amount of time allocated to instruction. Consequently, I assumed that the percentage of time allocated for instruction remained constant throughout the study. Therefore, for the purposes of this study, I assumed that field trips, assemblies, emergency drills, and various other non-instructional usages of time were the same for both the comparison and the experiential groups. Additionally, the study included an assumption that a similar proportion of allocated time resulted in student engagement with the content for both the comparison and the experiential groups.

A fourth limitation of this study was that the results may only pertain to the locale of this study and may not generalize to other regions. Additionally, the results of the quasi-experimental component may only pertain to Hispanic students and may not generalize to other populations. Similarly, since only one instructor was involved in this study, the results documented in this study may not be replicable by other instructors. Lastly, the results of this study may also be specific to the Calculus AB exam, and may not translate to other Advanced Placement exams.

Implication of the Results for Practice, Policy, and Theory

There has been an increase in the number of both male and female students taking the Calculus AB exam in the last fifteen years (College Board, 2017c). The increase in participation has not eliminated the gender achievement gap. While an increase in the quantity of instructional time produced statistically significant differences on the Calculus AB exam for Hispanic female students, it also produced improved results for Hispanic male students. It is unclear whether the strategy of increased instructional time would produce similar improvement in Calculus AB exam scores for students of all demographics, although it does seem like administrators should consider the strategy as possible means of combatting the gender achievement gap.

There has also been a sizable increase in the number Black and Hispanic students taking the Calculus AB exam over the course of the last 15 years, but this increase has not translated to higher Calculus AB exam scores for students of these ethnicities (College Board, 2017c). Consequently, Black and Hispanic students are less likely than their Asian and White counterparts are to attain scores on the Calculus AB exam that translate to college credit. According to OTL theory, schools have the ability to control three variables that may help improve the outcome for Black and Hispanic students, namely instructional time, content, and quality of instruction (Elliott & Bartlett, 2016).

Researchers have concluded that Black and Hispanic students are less likely to have teachers who possess advanced degrees, are certified, or are even qualified to teach high level math courses, but more likely to have inexperienced teachers, than their White peers (Corcoran & Evans, 2008; Darling-Hammond, 2010; Flores, 2007). No attempt was made in this study to determine whether an intentional plan to improve teacher quality, either by hiring teachers that were better qualified or by retaining and developing teachers, would be likely to increase student achievement for Black and Hispanic students on the Calculus AB exam. However, the discrepancy in experience, degrees in mathematics, and certification may not be a problem that school and district personnel, who most likely employ the best applicants they get, can easily solve.

In contrast, the principal focus of this study was whether increasing the amount of instructional time for Advanced Placement Calculus might shrink the size of the Calculus AB exam achievement gaps. Due to the annual fixed date for the Calculus AB exam, it is clear that districts that begin school after Labor Day are at a temporal disadvantage when compared to districts that begin as early as the beginning of August. A large percentage of Hispanic students are the children of recent immigrants. Students who are children of recent immigrants face familial obligations that are likely to be foreign to other students, such as suffering absences to provide translation services for parents or obligations to work long hours to help support the family. By scheduling an extra hour for Calculus practice at school, districts and schools can provide the essential support needed by students at home. However, administrators should be prepared to provide professional development for teachers to maximize the effects of the additional classroom time.

In conclusion, schools have the ability to control three critical variables that affect learning, namely instructional time, content, and quality of instruction (Elliott & Bartlett, 2016). However, College Board dictates the content for Advanced Placement Calculus, and improving the quality of instruction may require personnel changes or extensive amounts of professional development. However, the results of this study suggest that increasing the length of the instructional period may ameliorate the confluence of factors that negatively affect student achievement in Advanced Placement Calculus, such as gender, ethnicity, and socioeconomic status.

Recommendations for Further Research

This study began with an investigation of achievement gaps in the state of Michigan pertaining to the Calculus AB exam. Statistically significant differences were uncovered in each of the three segments of the achievement gap analysis. This meant that a gender achievement gap, an ethnic achievement gap, and a gender achievement gap for Hispanic participants was present on the Calculus AB exam. Several researchers have previously performed ethnic achievement comparisons of Advanced Placement scores between either Black or Hispanic students and White students. Some of the variance in Calculus AB exam scores seems to be attributable to gender and ethnicity (Holmes, 2013; Morris, 2013). However, Burney (2010) concludes that there are school level factors, such as school size, that account for some of the variance in student scores. Moreover, researchers have determined that the state of residency also is a factor in Advanced Placement exam scores (Davis, 2012; Koch, 2012; Wilson, 2013). Thus, a quantitative study that investigates the amount of variance attributable to ethnicity, gender, socioeconomic status, state of residency, and certain school level variables, such as school size, might help researchers better understand the phenomenon of achievement gaps pertaining to the Calculus AB exam.

While the gender and ethnic achievement gap analyses corroborate previous findings by researchers in various locales, the Hispanic gender analysis simply generates more questions for future researchers to answer. In example, three of those questions are: 1) Has a Hispanic gender gap in Calculus AB exam performance emerged where one did not previously seem to exist? 2) If a Hispanic gender achievement gap has emerged on the Calculus AB exam, then what spurred this development? 3) Why does gender now seem to be a stronger factor in Hispanic Calculus AB exam performance than in the rest of the population?

The second component of this study was a quasi-experiment to determine whether school and district personnel might redress the achievement gaps that exists between both the various ethnicities and the two genders by increasing the length of the instructional period. Statistically significant differences in score distribution existed in each of the four comparisons between the comparison group and the experiential group. Since Michigan has comparatively few Hispanic students that take the Calculus AB exam, researchers need to test the hypothesis that increasing the length of the yearlong instructional period for Advanced Placement Calculus from one hour to two hours may reduce the size of the achievement gaps in other locations and with larger sample sizes.

The conceptual framework for this study was OTL theory. The sole OTL variable examined within the course of this study was instructional time. However, the measurement of the variable of instructional time in this study was rudimentary. For example, the design of this study did not consider the amount of time lost to absences, snow days, assemblies, emergency drills, counseling, or school days after the Advanced Placement exam is taken. Boykin and Noguera (2011) recommend logging the amount of time that students spend performing mathematics; however, this sort of logging system might be overly burdensome on the instructor, the researcher, or both. However, documenting the days of student attendance might provide a vehicle to improve upon this quasi-experiment with matched samples based upon attendance and SAT group. Regardless of the manner in which the experiment is structured, future replication of this quasi-experiment should utilize a more sophisticated method of measuring the amount of instructional time.

As the quasi-experimental component in this study occurred in a mostly Hispanic school, it remains unclear whether the strategy of increased amount of instructional time might serve as a mechanism for providing Black students the OTL. This study took place in a Title I school; therefore, the subject school resided in a socioeconomically depressed zip code. Darling-Hammond (2010) laments that students of color in such schools do not have the same OTL that their peers in more affluent areas have. Consequently, this study should be replicated in schools that are predominantly Black, as well in schools that are socioeconomically depressed to determine whether this strategy may serve to redress the effects of urbanity and poverty, and in so doing provide students the opportunity to excel.

In this study, I compared achievement of all four ethnicities to determine whether there was a statistically significant difference in the distribution of scores for the four ethnicities. Absent from the literature is a direct comparison of Black students' Advanced Placement achievement to that of Hispanic students. Future research should compare the performance of Black and Hispanic students and subsequently investigate whether the strategy of increasing the amount of instructional time for Calculus yields similar benefits for Black and Hispanic students.

Conclusion

The twofold purpose of this study was to determine whether gender and ethnic achievement gaps exist on the Calculus AB exam in Michigan and to determine whether OTL theory can provide a way to reduce the size of achievement gaps on the Calculus AB exam. To measure the achievement gaps and test the aforementioned hypothesis, this study included an achievement gap analysis and a quasi-experimental component. The achievement gap analysis incorporated a gender achievement gap analysis, an ethnic achievement gap analysis, and a Hispanic gender achievement gap analysis. In the quasi-experimental component, the comparison group received one hour of daily Calculus instruction, while the experiential group received two hours of daily Calculus instruction. Grounded in OTL theory, there was an expectation that the experiential group would outperform the comparison group on the Calculus AB exam.

The results of the achievement gap analysis prompted the conclusion that there are both gender and ethnic achievement gaps on the Calculus AB exam in Michigan. The effect size of the gender achievement gap narrowed from .12 to .09 from 2002 to 2017, with a monotonically downward trend. However, the effect size for all four years was small. The effect size for the ethnic achievement gap has remained constant at .13 from 2007 to 2017, which implies that the ethnic achievement gap is not shrinking in size. The Hispanic gender achievement analysis was less conclusive, as a statistically significant difference in Calculus AB exam performance between Hispanic males and Hispanic females existed in only the most recent year of the study. The effect size for the Hispanic gender achievement gap analysis for 2017 was .19, which would seemingly indicate that gender is now a stronger correlate in Hispanic achievement on the Calculus AB exam than it is for the population as a whole. Thus, there are persistent gender and ethnic achievement gaps on the Calculus AB exam, and there may be an emergent Hispanic gender achievement gap.

The quasi-experimental component included four separate comparisons of the comparison group with the experiential group. All four comparisons yielded statistically significant differences in favor of the experiential group. The effect size for the comparison of all students was .48, which correlates to a medium effect (Bosco et al. 2015; Gignac & Szodorai, 2016). The effect size for only Hispanic student was .49, which also translates to a medium effect (Bosco et al. 2015; Gignac & Szodorai, 2016). The effect size for Hispanic males was .63, which is interpreted as a large effect, while the effect size for Hispanic females was .45, a medium effect size (Bosco et al. 2015; Gignac & Szodorai, 2016). These effect sizes support the conclusion that an increase in the amount of time invested in instruction can result in higher student achievement on the Calculus AB exam.

Clearly, the achievement gaps pertaining to the Calculus AB exam in Michigan still exist. Both gender and ethnic achievement gaps remain. The quasi-experimental component of this study, an attempt to test a potential solution to the problems of the Calculus AB achievement gaps, featured a predominantly Hispanic population. The large percentage of Hispanic participants in this study may signify that the results of this study may not generalize to other populations. However, the increase in achievement for the experiential group, in light of OTL theory, inspires the conclusion that schools may be able to increase achievement on the Calculus AB exam by increasing the length of the instructional period.

References

Abedi, J., Courtney, M., Leon, S., Kao, J., & Azzam, T. (2006). *English language learners and math achievement: A study of Opportunity to Learn and language accommodation* (ED495848). National Center for Research on Evaluation, Standards, and Student Testing. https://eric.ed.gov/?id=ED495848

Abedi, J., & Herman, J. (2010). Assessing English language learners' opportunity to learn mathematics: Issues and limitations. *Teachers College Record*, *112*(3), 723–746. http://www.tcrecord.org.cupdx.idm.oclc.org/library/content.asp?contentid=15878

Acemoglu, D., & Autor, D. (2012). What does human capital do? A review of Goldin and Katz's the race between education and technology. *Journal of Economic Literature*, *50*(2), 426–463. https://doi.org/10.1257/jel.50.2.426

Ackerman, P. L., Kanfer, R., & Calderwood, C. (2013). High school Advanced Placement and student performance in college: STEM majors, non-STEM majors, and gender differences. *Teachers College Record*, *115*(10), 1–43. http://www.tcrecord.org.cupdx.idm.oclc.org/library/content.asp?contentid=17149

ACT, Inc. (2014). *The ACT® Technical Manual*. https://www.act.org/content/dam/act/unsecured/documents/ACT_Technical_Manual.pdf

Adams, K. A., & Lawrence, E. K. (2015). *Research methods, statistics, and applications*. Thousand Oaks, CA: SAGE.

Aguirre-Muñoz, Z., & Amabisca, A. A. (2010). Defining opportunity to learn for English language learners: Linguistic and cultural dimensions of ELLs' instructional contexts. *Journal of Education for Students Placed at Risk*, *15*(3), 259–278. https://doi.org/10.1080/10824669.2010.495691

Ahmad, S. N., Mahadi, S., Yusri, M. Y., Yusop, H., Ali, M. N., & Heng, C. H. (2017). Factors related to students' performance in Calculus. *Journal of Applied Environmental and Biological Sciences*, *7*(6), 51–56.

Albalawi, A. S. (2018). The effect of using flipped classroom in teaching Calculus on students' achievements at University of Tabuk. *International Journal of Research in Education and Science*, *4*(1), 198–207. https://doi.org/10.21890/ijres.383137

Alford-Stephens, T. (2016). *Differences in mathematics skills of Texas high school boys as a function of ethnicity/race and economic status: A multiyear statewide study* (Doctoral dissertation). https://shsu-ir.tdl.org/shsu-ir/handle/20.500.11875/39

Altonji, J., Blom, E., & Meghir, C. (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annual Review of Economics*, *4*(1), 185–223. https://doi.org/10.3386/w17985

American Psychological Association. (2010). *APA Manual (Publication manual of the American Psychological Association)* (6th ed.). Washington, DC: Author.

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science*, *344*(6186), 843–851. https://doi.org/10.1126/science.1251868

Ben-Zeev, A., Paluy, Y., Milless, K., Goldstein, E., Wallace, L., Marquez-Magana, L., & Estrada, M. (2017). 'Speaking truth' protects underrepresented minorities' intellectual performance and safety in STEM. *Education Scienc*es, *7*(2), 65. https://doi.org/10.3390/educsci7020065

Bauer, E. (2016). Practising kinship care: Children as language brokers in migrant families. *Childhood*, *23*(1), 22–36. https://doi.org/10.1177/0907568215574917

Bolaños-Lopez, J. F. (2017). Unauthorized immigration in the United States: Impact, media representation and legalization reform. *INNOVA Research Journal*, 2(8), 281–293. http://www.journaluidegye.com/magazine/index.php/innova/

Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology*, *100*(2), 431–449. https://doi.org/10.1037/a0038047

Boykin, A. W., & Noguera, P. (2011). *Creating the opportunity to learn: Moving from research to practice to close the achievement gap*. Alexandria, VA: ASCD.

Buckley, L. A. (2010). Unfulfilled hopes in education for equity: redesigning the mathematics curriculum in a US high school. *Journal of Curriculum Studies*, *42*(1), 51–78. https://doi.org/10.1080/00220270903148065

Buford, B. J. (2012). *A mixed methods study: African American students' performance trends and perceptions towards Advanced Placement literature courses and examinations* (Doctoral dissertation). ProQuest Dissertations and Theses database. (1318871587)

Bumpous, D. A. (2015). *The relationship between Advanced Placement participation and college graduation rates* (Doctoral dissertation). ProQuest Dissertations and Theses database. (10131986)

Burney, V. H. (2010). High achievement on Advanced Placement exams: The relationship of school-level contextual factors to performance. *The Gifted Child Quarterly*, *54*(2), 116–126. https://doi.org/10.1177/0016986209355972

Carroll, J. B. (1989). The Carroll Model: A 25-year retrospective and prospective view. *Educational Researcher*, *18*(1), 26–31. https://doi.org/10.2307/1176007

Cashmere, D. (2018). Teaching high school cultural studies in the age of Trump. *Radical Teacher*, *111*, 35–45. https://doi.org/10.5195/rt.2018.515

Central Michigan University. (2015). Academic policies and procedures. https://www.cmich.edu/ess/registrar/Documents/records/apcredit.pdf

Chajewski, M., Mattern, K. D., & Shaw, E. J. (2011). Examining the role of Advanced Placement® exam participation in 4-year college enrollment. *Educational Measurement: Issues and Practice*, *30*(4), 16–27. https://doi.org/10.1111/j.1745-3992.2011.00219.x

Champion, J., & Mesa, V. (2016). Factors affecting calculus completion among U.S. high school students. In D. M. Bressoud (Ed.), *The role of calculus in the transition from high school to college mathematics* (pp. 9–26). Washington, DC: The Mathematical Association of America and the National Council of Teachers of Mathematics.

Cimpian, J. R., Lubienski, S. T., Timmer, J. D., Makowski, M. B., & Miller, E. K. (2016). Have gender gaps in math closed? Achievement, teacher perceptions, and learning behaviors across two ECLS-K cohorts. *AERA Open*, *2*(4), 1–19. https://doi.org/10.1177/2332858416673617

Cisneros, J., Gomez, L. M., Powers, J. M., Holloway-Libell, J., & Corley, K. M. (2014). The Advanced Placement opportunity gap in Arizona: Access, participation, and success. *AASA Journal of Scholarship and Practice*, *11*(2), 20–33.

Claudet, J. G. (2016). Deconstructing the breakthrough leadership thinking of visionary social change agents—Insights and strategies for leading transformative change from four case studies. *Advances in Applied Sociology*, *6*(9), 271–318. https://doi.org/aasoci.2016.69023

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in high school. *Journal of Human Resources*, *45*(3), 655–681. https://doi.org/10.3368/jhr.45.3.655

Cobb, F., & Russell, N. M. (2015). Meritocracy or complexity: problematizing racial disparities in mathematics assessment within the context of curricular structures, practices, and discourse. *Journal of Education Policy*, *30*(5), 631–649. https://doi.org/10.1080/02680939.2014.983551

College Board. (2002). Michigan summary- 2002. http://media.collegeboard.com/digitalServices/pdf/research/mi_2002.xls

College Board. (2012). *Calculus: Calculus AB, Calculus BC course description*. http://apcentral.collegeboard.com/apc/public/repository/ap-calculus-course-description.pdf

College Board. (2014). *The 10th annual AP report to the nation*. http://media.collegeboard.com/digitalServices/pdf/ap/rtn/10th-annual/10th-annual-ap-report-to-the-nation-single-page.pdf

College Board. (2015a). *Annual AP program participation 1956–2015*. https://secure-media.collegeboard.org/digitalServices/pdf/research/2015/2015-Annual-Participation.pdfd.org/digitalServices/pdf/research/2015/2015-Annual-Participation.pdf

College Board. (2015b). *AP exam volume changes (2005–2015)*. https://secure-media.collegeboard.org/digitalServices/pdf/research/2015/2015-Exam-Volume-Change.pdfmedia.collegeboard.org/digitalServices/pdf/research/2015/2015-Exam-Volume-Change.pdf

College Board. (2015c). *SAT: Comparing SAT and ACT scores*. https://secure-media.collegeboard.org/digitalServices/pdf/sat/cSAT-concordance-flyer.pdf

College Board. (2015d). *Test specifications for the redesigned SAT*. https://collegereadiness.collegeboard.org/pdf/test-specifications-redesigned-sat-1.pdf

College Board. (2016a). *Course and exam description: AP Calculus AB and AP Calculus BC, including the curriculum framework*. New York, NY.

College Board. (2016b). K–12 educator brief: Concordance. https://collegereadiness.collegeboard.org/pdf/k-12-educator-brief-sat-concordance.pdf

College Board. (2017a). AP central- Past exam questions for Calculus AB. http://apcentral.collegeboard.com/apc/members/exam/exam_information/232050.html

College Board. (2017b). *Bulletin for AP students and parents.* https://apcentral.collegeboard.org/pdf/ap-student-bulletin-2017-18.pdf

College Board. (2017c). Michigan summary- 2017. http://media.collegeboard.com/digitalServices/pdf/research/2017/Michigan_Summary_17.xls

Conley, D. T., & French, E. M. (2013). Student ownership of learning as a key component of college readiness. *American Behavioral Scientist*, *58*(8), 1018–1034. https://doi.org/10.1177/0002764213515232

Corcoran, S. P., & Evans, W. N. (2008). The role of inequality in teacher quality. In K. Magnuson & J. Waldfogel (Eds.), *Steady gains and stalled progress: Inequality and the Black-White test score gap* (pp. 212–249). New York, NY: Russell Sage Foundation.

Crawford, C., & Cribb, J. (2013). *Reading and maths skills at age 10 and earnings in later life: A brief analysis using the British Cohort Study* (3). Centre for Analysis of Youth Transitions. https://www.ifs.org.uk/caytpubs/CAYTreport03.pdf

Darling-Hammond, L. (2010). *The flat world and education: How America's commitment to equity will determine our future*. New York, NY: Teachers College Press.

Davis, C. M. (2012). *A multi-year comparison of Advanced Placement scores for Black students in Texas, New York, and Florida* (Doctoral dissertation). ProQuest Dissertations and Theses database. (3571391)

Davis, C. M., Slate, J. R., Moore, G. W., & Barnes, W. (2014). Black student performance on Advanced Placement exams: A multiyear, multistate comparison. *The Online Journal of New Horizons in Education*, *4*(2), 21–36.

Dorans, N. J. (2004). Equating, concordance, and expectation. *Applied Psychological Measurement*, *28*(4), 227–246. https://doi.org/10.1177/0146621604265031

Drost, E. A. (2011). Validity and reliability in social science research. *Education Research and Perspectives*, *38*(1), 105–123.

Easey, M., & Gleeson, J. (2016). The relevance of mathematics: Leaders and teachers as gatekeeper for Queensland senior calculus mathematics. In B. White, M. Chinnappan, & S. Trenholm (Eds.), *Opening up mathematics education research (Proceedings of the 39th annual conference of the Mathematics Education Research Group of Australasia)* (pp. 198–205). Adelaide, Australia: Mathematics Education Research Group of Australasia.

Elliott, S. N., & Bartlett, B. J. (2016). Opportunity to learn. *Oxford Handbooks Online.* https://doi.org/10.1093/oxfordhb/9780199935291.013.70

Ellison, G., & Swanson, A. (2010). The gender gap in secondary school mathematics at high achievement levels: Evidence from the American Mathematics Competitions. *Journal of Economic Perspectives*, *24*(2), 109–128. https://doi.org/10.1257/jep.24.2.109

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, *136*(1), 103–127. https://doi.org/10.1037/a0018053

Escalante, J., & Dirmann, J. (1990). The Jaime Escalante Math Program. *The Journal of Negro Education*, *59*(3), 407. https://doi.org/10.2307/2295573

Fagerland, M. W., Laake, P., & Lydersen, S. (2017). *Statistical analysis of contingency tables*. Boca Raton, FL: CRC Press.

Ferris State University. (2016). Advanced Placement: Course equivalent and credits. https://ferris.edu/admissions/testing/docs/AP-Course-Equivalents-and-Courses.pdf

Floden, R. E. (2002). The measurement of Opportunity to Learn. In A. C. Porter & A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 231–266). Washington, DC: National Academy Press.

Flores, A. (2007). Examining disparities in mathematics education: Achievement gap or opportunity gap? *The High School Journal*, *91*(1), 29–42. https://doi.org/10.1353/hsj.2007.0022

Fryer, R. G., & Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics*, *2*(2), 210–240. https://doi.org/10.1257/app.2.2.210

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069

Gilroy, M. (2006). Jaime Escalante still stands and delivers: A 30-year overnight success. *The Hispanic Outlook in Higher Education*, *16*(11), 35.

Glennon, L., & Mohler, M. (1999). *Those who can-- teach!: Celebrating teachers who make a difference*. Berkeley, CA: Wildcat Canyon Press.

Gonzales, A. (2017). Trumpism, authoritarian neoliberalism, and subaltern Latina/o politics. *Aztlán: A Journal of Chicano Studies*, *42*(2), 147–164. http://www.chicano.ucla.edu/publications/aztl%C3%A1n-journal-chicano-studies-0

Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology*, *29*(1), 17–28. https://doi.org/10.1016/j.appdev.2007.10.004

Goodman, S. (2016). Learning from the Test: Raising Selective College Enrollment by Providing Information. *The Review of Economics and Statistics*, *98*(4), 671–684. https://doi.org/10.1162/REST_a_00600

Gradillas, H., & Jesness, J. (2010). *Standing and delivering: What the movie didn't tell*. Lanham, MD: Rowman & Littlefield Education.

Gravetter, F. J., & Wallnau, L. B. (2009). *Statistics for the behavioral sciences* (8th ed.). Belmont, CA: Cengage Learning.

Guan, S. A., Greenfield, P. M., & Orellana, M. F. (2014). Translating Into Understanding. *Journal of Adolescent Research*, *29*(3), 331-355. https://doi.org/10.1177/0743558413520223

Gullatt, D. E. (2006). Block scheduling: The effects on curriculum and student productivity. *NASSP Bulletin*, *90*(3), 250–266. https://doi.org/10.1177/0192636506292382

Guskey, T. R., & Kifer, E. (1995). *Evaluation of a high school block schedule restructuring program* (ED384652). http://files.eric.ed.gov/fulltext/ED384652.pdf

Handwerk, P., Tognatta, N., Coley, R. J., & Gitomer, D. H. (2008). *Access to success: Patterns of Advanced Placement participation in U.S. high schools*. Princeton, NJ: Educational Testing Service.

Hansen, D., Gutman, M., & Smith, J. (2000). Scheduling AP classes in a 2X4 block schedule: The Mayfield plan. *Phi Delta Kappan*, *82*(3), 209–211. https://doi.org/10.1177/003172170008200305

Hattie, J. A. (2008). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.

Henry, B. B. (2008). *Factors influencing student achievement in Advanced Placement and Honors courses* (3324979) (Doctoral dissertation). ProQuest Dissertations and Theses database. (UMI No. 3324979)

Hernandez, M. (2014). The relationships between mathematics achievement and socioeconomic status. *JEP: Ejournal of Education Policy*, 1–9.

Holmes, M. A. (2013). *Ethnic and gender differences in Advanced Placement exam performance: A multiyear national analysis* (Doctoral dissertation). ProQuest Dissertations and Theses database. (3571401)

Hyde, J. S. (2016). Sex and cognition: gender and cognitive functions. Current Opinion in Neurobiology, 38, 53–56. https://doi.org/10.1016/j.conb.2016.02.007

Inzlicht, M., & Kang, S. K. (2010). Stereotype threat spillover: How coping with threats to social identity affects aggression, eating, decision making, and attention. *Journal of Personality and Social Psychology*, *99*(3), 467–481. https://doi.org/10.1037/a0018951

Inzlicht, M., & Schmader, T. (2012). Introduction. In M. Inzlicht & T. Schmader (Eds.), *Stereotype threat: Theory, process, and application* (pp. 3–16). New York, NY: Oxford University Press.

James, J. (2013). *The surprising impact of high school math on job market outcomes* (2013–14). Cleveland, OH: Federal Reserve Bank of Cleveland.

James, M. C., Butterfield, V., Jones, K. L., & Mokuria, V. G. (2017). Opportunity for all: An analysis of equality and equity in Advanced Placement programming in a US high school. *International Journal of Innovation and Research in Educational Sciences*, *4*(1), 15–21. http://www.ijires.org/

Jara, T. D. (2013). *Hispanic student performance on Advanced Placement Exams: A multiyear, national investigation* (Doctoral dissertation). ProQuest Dissertations & Theses Global database. (UMI No. 3571400)

JBHE Foundation. (2004). College-bound Black students are making inroads in Advanced Placement tests. *The Journal of Blacks in Higher Education*, (44), 99. https://doi.org/10.2307/4133753

Johnson, R. B., & Christensen, L. (2013). *Educational research: Quantitative, qualitative, and mixed approaches* (5th ed.). Thousand Oaks, CA: SAGE.

Judson, E. (2017). Science and mathematics Advanced Placement exams: Growth and achievement over time. *The Journal of Educational Research*, *110*(2), 209–217. https://doi.org/10.1080/00220671.2015.1075188

Judson, E., & Hobson, A. (2015). Growth and achievement trends of Advanced Placement (AP) exams in American high schools. *American Secondary Education*, *43*(2), 59–76.

Julie, C., Mbekwa, M., & Simon, M. (2016). The effect size of an intervention focusing on the use of previous national senior certificate mathematics examination papers. *Journal of Educational Studies*, *15*(1), 1–22. http://journals.co.za/content/journal/jeds

Kalamazoo College. (2017). Math department: Advanced Placements credits. Kalamazoo College. https://reason.kzoo.edu/math/ap_credits/

Kapitanoff, S., & Pandey, C. (2017). Stereotype threat, anxiety, instructor gender, and underperformance in women. *Active Learning in Higher Education*, *18*(3), 213–229. https://doi.org/10.1177/1469787417715202

Katz, V. (2014). Children as brokers of their immigrant families' health-care connections. *Social Problems*, *61*(2), 194–215. https://doi.org/10.1525/sp.2014.12026

Keen, C. M. (1996). *An investigation of the achievement of 4x4 block-scheduled Advanced Placement Calculus AB students as indicated by their scores on the 1995 Advanced Placement Calculus AB examination* (Doctoral dissertation). ProQuest Dissertations and Theses database. (9717679)

Kellow, J. T., & Jones, B. D. (2008). The effects of stereotypes on the achievement gap: Reexamining the academic performance of African American high school students. *Journal of Black Psychology*, *34*(1), 94–120. https://doi.org/10.1177/0095798407310537

Keng, L., & Dodd, B. G. (2008). *A comparison of college performances of AP and non-AP student groups in 10 subject areas* (2008–7). New York, NY: College Board.

Kester, Donald. (1993). *An instructional guide concerning the highly successful teaching & motivating practices of Jaime Escalante for the Escalante Math Project at East Los*

Angeles College (ED358173). Los Angeles County Office of Education. https://eric.ed.gov/?id=ED358173

Kettler, T., & Hurst, L. T. (2017). Advanced academic participation: A longitudinal analysis of ethnicity gaps in suburban schools. *Journal for the Education of the Gifted*, *40*(1), 3–19. https://doi.org/10.1177/0162353216686217

Kidron, Y., & Lindsay, J. (2014). *The effects of increased learning time on student academic and nonacademic outcomes: Findings from a meta-analytic review* (REL 2014–015). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Appalachia.

Kim, C., Tamborini, C. R., & Sakamoto, A. (2015). Field of study in college and lifetime earnings in the United States. Sociology of Education, 88(4), 320–339. https://doi.org/10.1177/0038040715602132

Kitsantas, A., Cheema, J., & Ware, H. W. (2011). Mathematics achievement: The role of homework and self-efficacy beliefs. *Journal of Advanced Academics*, *22*(2), 310–339. https://doi.org/10.1177/1932202x1102200206

Klopfenstein, K. (2004a). Advanced Placement: Do minorities have equal opportunity? *Economics of Education Review*, *23*(2), 115–131. https://doi.org/10.1016/s0272-7757(03)00076-1

Klopfenstein, K. (2004b). The Advanced Placement expansion of the 1990s: How did traditionally underserved students fare? *Education Policy Analysis Archives*, *12*(68), 1–13. https://doi.org/10.14507/epaa.v12n68.2004

Knudson-Martin, J. (2011). A Combined Model for Understanding Motivation. American *International Journal of Contemporary Research*, *1*(2), 11–16. https://doi.org/10.30845/aijcr

Koch, B. M. (2012). *A comparison of advanced placement scores for Hispanic students from California, Texas, and Arizona*. ProQuest Dissertations & Theses Global. (UMI No. 1157543993)

Koch, B., Slate, J. R., & Moore, G. W. (2014). Advanced placement mathematics: A comparison of scores for Hispanic students from California, Texas, and Arizona. *Jeppa*, *3*(2), 1–27.

Kodippili, A. (2011). Parents' education level in students' mathematics achievement; Do school factors matter? *Academic Leadership: The Online Journal*, *9*(1), Article 39.

Kolluri, S. (2018). Advanced Placement: The Dual Challenge of Equal Access and Effectiveness. *Review of Educational Research*, 1–41. https://doi.org/10.3102/0034654318787268

Lamb, S., & Fullarton, S. (2002). Classroom and school factors affecting mathematics achievement: A comparative study of Australia and the United States using TIMSS. *Australian Journal of Education*, *46*(2), 154–171. https://doi.org/10.1177/000494410204600205

Lewis, S., Simon, C., Uzzell, R., Horwitz, A., & Casserly, M. (2010). *A call for change: The social and educational factors contributing to the outcomes of Black males in urban schools* (ED512669). Council of the Great City Schools. https://eric.ed.gov/?id=ED512669

Lightner, L. (2009). *The Effects of block scheduling on AP Calculus AB student achievement* (Master's thesis, Bemidji State University).

http://faculty.bemidjistate.edu/grichgels/MastersPapers/Larry%20Lightner%20Final%20Paper.pdf

Mahtani, K., Spencer, E. A., Brassey, J., & Heneghan, C. (2018). Catalogue of bias: observer bias. *BMJ Evidence-Based Medicine*, *23*(1), 23–24. https://doi.org/10.1136/ebmed-2017-110884

Marzano, R. J., Pickering, D., & Pollock, J. E. (2001). *Classroom management that works: Research-based strategies for every teacher*. Alexandria, VA: ASCD.

Mathews, J. (1988). *Escalante: The best teacher in America*. New York, NY: Henry Holt.

Mattern, K. D., Marini, J. P., & Shaw, E. J. (2013). *Are AP students more likely to graduate from college on time?* (2013–5) New York, NY: College Board.

Melsom, K. D. (2008). *A longitudinal study of gender differences in Advanced Placement participation and scoring* (Doctoral dissertation). ProQuest Dissertations and Theses database. (3329836)

Michigan Department of Education. (2018). MI school data: Student counts. https://www.mischooldata.org/DistrictSchoolProfiles2/StudentInformation/StudentCounts/StudentCount.aspx

Michigan State University. (2017). College Board Advanced Placement program equivalencies. https://admissions.msu.edu/documents/MSU_AP_Equivalencies.pdf

Moller, S., Stearns, E., Southworth, S., & Potochnick, S. (2013). Changing course: the gender gap in college selectivity and opportunities to learn in the high school curriculum. *Gender and Education*, *25*(7), 851–871. https://doi.org/10.1080/09540253.2013.853028

Moore, G. W., Combs, J. P., & Slate, J. R. (2012). Advanced Placement exams participation and performance: A national study of gender differences. *E-International Journal of Educational Research*, *3*(3), 18–32.

Moore, G. W., & Slate, J. R. (2008). Who's taking the Advanced Placement courses and how are they doing: A statewide two-year study. *The High School Journal*, *92*(1), 56–67. https://doi.org/10.1353/hsj.0.0013

Moore, G., & Slate, J. (2010). Advanced Placement course and American Indian performance. *American Secondary Education*, *38*(2), 73–94.

Moore, G., & Slate, J. R. (2011). A multi-year analysis of Asian gender differences on Advanced Placement exams. *International Journal of Educational Leadership Preparation*, *6*(4), 1–13.

Morgan, R., & Klaric, J. (2007*). AP students in college: An analysis of five-year academic careers* (2007–4). New York, NY: College Board.

Morris, J. B. (2013). *Gender differences in STEM related Advanced Placement exams* (Doctoral dissertation). ProQuest Dissertations and Theses database. (UMI No. 3571393)

Morris, J., & Slate, J. R. (2012). Advanced Placement math scores from 2000 to 2010: Does gender still matter? *E-International Journal of Educational Research*, *3*(1), 33–48.

Morris, J. B., Slate, J. R., & Moore, G. W. (2015). Gender and STEM related Advanced Placement exams: A review of the literature. *The Online Journal of New Horizons in Education*, *5*(1), 55–71.

Niederle, M., & Vesterlund, L. (2010). Explaining the gender gap in math test scores: The role of competition. *Journal of Economic Perspectives*, *24*(2), 129–144. https://doi.org/10.1257/jep.24.2.129

Obama, B. H. (2016, October 17). Remarks by the president on education. https://obamawhitehouse.archives.gov/the-press-office/2016/10/17/remarks-president-education

Onwuegbuzie, A. J. (2000). *Expanding the framework of internal and external validity in quantitative research* (ED 448205). https://eric.ed.gov/?id=ED448205

Ottmar, E. R., Decker, L. E., Cameron, C. E., Curby, T. W., & Rimm-Kaufman, S. E. (2013). Classroom instructional quality, exposure to mathematics instruction and mathematics achievement in fifth grade. *Learning Environments Research*, *17*(2), 243–262. https://doi.org/10.1007/s10984-013-9146-6

Pennington, C. R., & Heim, D. (2016). Creating a critical mass eliminates the effects of stereotype threat on women's mathematical performance. *British Journal of Educational Psychology*, *86*(3), 353–368. https://doi.org/10.1111/bjep.12110

Perry, M. J. (2016). 2016 SAT test results confirm pattern that's persisted for 50 years- High school boys are better at math than girls. http://www.aei.org/publication/2016-sat-test-results-confirm-pattern-thats-persisted-for-45-years-high-school-boys-are-better-at-math-than-girls/

Peterson, B. L. (1989). *Factors affecting achievement in the first course in calculus at the high school level* (9017233) (Doctoral dissertation). ProQuest Dissertations and Theses database. (Order No. 9017233)

Reardon, S. F., Greenberg, E., Kalogrides, D., Shores, K. A., & Valentino, R. A. (2012). *Trends in academic achievement gaps in the era of No Child Left Behind* (ED530121). Society for Research on Educational Effectiveness. https://eric.ed.gov/?id=ED530121

Remler, D. K., & Van, R. G. (2014). *Research methods in practice: Strategies for description and causation* (2nd ed.). Thousand Oaks, CA: SAGE.

Rowland, J. (2014). *Number of instructional days/hours in the school year* (ED561932). Education Commission of the States. https://eric.ed.gov/?id=ED561932

Rubin, A. (2010). *Statistics for evidence-based practice and evaluation* (2nd ed.). Belmont, CA: Brooks / Cole.

Sadler, P. M., & Sonnert, G. (2016). Factors influencing success in introductory college Calculus. In D. M. Bressoud (Ed.), *The role of calculus in the transition from high school to college mathematics* (pp. 53–66). Washington, DC: Mathematical Association of America and the National Council of Teachers of Mathematics.

Sahin, A., Cavlazoglu, B., & Zeytuncu, Y. E. (2015). Flipping a college Calculus course: A case study. *Educational Technology & Society*, *18*(3), 142–152.

Sánchez, P. (2014). Dignifying every day: Policies and practices that impact immigrant students. *Language Arts*, *91*(5), 363–371. http://www2.ncte.org/resources/journals/language-arts/

Shaw, E. J., Marini, J. P., & Mattern, K. D. (2012). Exploring the utility of Advanced Placement participation and performance in college admission decisions. *Educational and Psychological Measurement*, *73*(2), 229–253. https://doi.org/10.1177/0013164412454291

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., … Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, *23*(7), 691–697.

Silva, C. M. (2010). Ex post facto study. In N. J. Salkind (Ed.), *Encyclopedia of research design: Volume 3* (pp. 465–466). https://doi.org/10.4135/9781412961288.n145

Smith, R., & Camara, W. (1998). *Block schedules and student performance on AP examinations* (RN-03). New York, NY: College Board.

Smith, J., Hurwitz, M., & Avery, C. (2017). Giving college credit where it is due: Advanced Placement exam scores and college outcomes. *Journal of Labor Economics*, *35*(1), 67–147. https://doi.org/10.1086/687568

Solórzano, D. G., & Ornelas, A. (2004). A critical race analysis of Latina/o and African American Advanced Placement enrollment in public high schools. *The High School Journal*, *87*(3), 15–26. https://doi.org/10.1353/hsj.2004.0003

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811. https://doi.org/10.1037//0022-3514.69.5.797

Stevens, F. I. (1996). *The need to expand the opportunity to learn conceptual framework: Should students, parents, and school resources be included?* (ED397523). American Educational Research Association. https://eric.ed.gov/?id=ED397523

Stricker, L. J. (1998). *Inquiring about examinees' ethnicity and sex: Effects on AP Calculus AB examination performance* (98–1). New York, NY: College Entrance Examination Board.

Stricker, L. J., & Ward, W. C. (2004). Stereotype threat, inquiring about test takers' ethnicity and gender, and standardized test performance. *Journal of Applied Social Psychology*, *34*(4), 665–693. https://doi.org/10.1111/j.1559-1816.2004.tb02564.x

Tosto, M. G., Asbury, K., Mazzocco, M. M., Petrill, S. A., & Kovas, Y. (2016). From classroom environment to mathematics achievement: The mediating role of self-perceived ability and subject interest. *Learning and Individual Differences*, *50*, 260–269.

Underwood, A. S. (2014). *A comparative study of the effect of block scheduling and traditional scheduling on student achievement for the Florida Algebra 1 End-of-Course Examination* (Doctoral dissertation). http://etd.fcla.edu/CF/CFE0005433/underwood_final_pdf.pdf

U.S. Department of Education. (2012). Definitions. https://www.ed.gov/race-top/district-competition/definitions

University of Michigan. (2017). AP, IB credit: Undergraduate admissions. https://admissions.umich.edu/apply/freshmen-applicants/ap-ib-credit

University of Michigan-Dearborn. (2017). Advanced Placement (AP) credit. https://umich.app.box.com/v/advanced-placement-2017-2018

Urrieta, L. (2016). Diasporic community smartness: Saberes (knowings) beyond schooling and borders. *Race Ethnicity and Education*, *19*(6), 1186–1199. https://doi.org/10.1080/13613324.2016.1168541

VanSciver, J. H. (2006). Closing the diversity gap in Advanced Placement course enrollment. *Multicultural Perspectives*, *8*(3), 56–58. https://doi.org/10.1207/s15327892mcp0803_10

Walker, S. A., & Pearsall, L. D. (2012). Barriers to Advanced Placement for Latino students at the high-school level. *Roeper Review*, *34*(1), 12–25. https://doi.org/10.1080/02783193.2012.627549

Walkowiak, T. A., Pinter, H. H., & Berry, R. Q. (2017). A reconceptualized framework for 'Opportunity to Learn' in school mathematics. *Journal of Mathematics Education at Teachers College*, *8*(1), 7–18. http://journals.tc-library.org/index.php/matheducation/article/view/1210/737

Wallace, C. (2013). *Comparing traditional period and semester block in high school mathematics: Effect on algebra I end-of-course assessment* (Doctoral dissertation). ProQuest Dissertations and Theses database. (3557706)

Walsh, K. S. (2011*). Schedule structure and student performance: Through the view of advanced placement* (Doctoral dissertation). ProQuest Dissertations and Theses database. (3459167)

Wayne State University. (2015). Advanced Placement policy for awarding credit. http://transfercredit.wayne.edu/pdf/appolicy20151117.pdf

Wijaya, A., Van den Heuvel-Panhuizen, M., & Doorman, M. (2015). Opportunity-to-learn context-based tasks provided by mathematics textbooks. *Educational Studies in Mathematics*, *89*(1), 41–65. https://doi.org/10.1007/s10649-015-9595-1

Wijaya, A., Van den Heuvel-Panhuizen, M., Doorman, M., & Veldhuis, M. (2018). Opportunity-to-Learn to solve context-based mathematics tasks and students' performance in solving these tasks – Lessons from Indonesia. *Eurasia Journal of Mathematics, Science and Technology Education*, *14*(10), 1–20. https://doi.org/10.29333/ejmste/93420

Wilson, J. L. (2013). *An examination of advanced placement scores for Black male students from Connecticut, Florida, Maryland, Massachusetts, and Texas* (Doctoral dissertation). ProQuest Dissertations & Theses Global. (UMI No. 3579553)

Witt, P. A., Ulmer, J. D., Burris, S., Brashears, T., & Burley, H. (2014). A Comparison of Student Engaged Time in Agriculture Instruction. *Journal of Agricultural Education*, *55*(2), 16–32. https://doi.org 10.5032/jae.2014.02016

Woods, J. R. (2015). *Testing trends: Considerations for choosing and using assessments* (ED561807). Education Commission of the States. https://eric.ed.gov/?id=ED561807

Zelkowski, J. (2010). Secondary mathematics: Four credits, block schedules, continuous enrollment? What maximizes college readiness? *The Mathematics Educator*, *20*(1), 8–21.

Zepeda, S. J., & Mayers, R. S. (2006). An analysis of research on block scheduling. *Review of Educational Research*, *76*(1), 137–170. https://doi.org/10.3102/00346543076001137

Appendix A: Statement of Original Work

The Concordia University Doctorate of Education Program is a collaborative community of scholar-practitioners, who seek to transform society by pursuing ethically-informed, rigorously-researched, inquiry-based projects that benefit professional, institutional, and local educational contexts. Each member of the community affirms throughout their program of study, adherence to the principles and standards outlined in the Concordia University Academic Integrity Policy. This policy states the following:

Statement of academic integrity.

As a member of the Concordia University community, I will neither engage in fraudulent or unauthorized behaviors in the presentation and completion of my work, nor will I provide unauthorized assistance to others.

Explanations:

What does "fraudulent" mean?

"Fraudulent" work is any material submitted for evaluation that is falsely or improperly presented as one's own. This includes, but is not limited to texts, graphics and other multi-media files appropriated from any source, including another individual, that are intentionally presented as all or part of a candidate's final work without full and complete documentation.

What is "unauthorized" assistance?

"Unauthorized assistance" refers to any support candidates solicit in the completion of their work, that has not been either explicitly specified as appropriate by the instructor, or any assistance that is understood in the class context as inappropriate. This can include, but is not limited to:

- Use of unauthorized notes or another’s work during an online test
- Use of unauthorized notes or personal assistance in an online exam setting
- Inappropriate collaboration in preparation and/or completion of a project
- Unauthorized solicitation of professional resources for the completion of the work.

Statement of Original Work

I attest that:

1. I have read, understood, and complied with all aspects of the Concordia University Portland Academic Integrity Policy during the development and writing of this dissertation.

2. Where information and/or materials from outside sources has been used in the production of this dissertation, all information and/or materials from outside sources has been properly referenced and all permissions required for use of the information and/or materials have been obtained, in accordance with research standards outlined in the Publication Manual of The American Psychological Association

Clint Prong

Digital Signature

Clint Prong

Name (Typed)

June 12, 2018

Date