2022

CHALLENGE BY CHOICE AND ITS IMPACT ON ENROLLMENT AND ACHIEVEMENT FOR ADVANCED PLACEMENT STUDENTS

Jessica Williams

CHALLENGE BY CHOICE AND ITS IMPACT ON ENROLLMENT AND ACHIEVEMENT FOR ADVANCED PLACEMENT STUDENTS

A dissertation submitted in partial fulfillment

of the requirements for the degree of

DOCTOR OF EDUCATION

to the faculty of the

DEPARTMENT OF ADMINISTRATIVE AND INSTRUCTIONAL LEADERSHIP

of

THE SCHOOL OF EDUCATION

at

ST. JOHN'S UNIVERSITY

New York

by

Jessica Williams

Date Submitted 3/8/2022 Date Approved 5/17/2022

_____	_____
Jessica Williams	Dr. Erin Fahle

ABSTRACT

CHALLENGE BY CHOICE AND ITS IMPACT ON ENROLLMENT AND ACHIEVEMENT FOR ADVANCED PLACEMENT STUDENTS

Jessica Williams

The purpose of this study was to determine the impact of a Challenge by Choice policy instituted in 2011 on AP course taking and performance in a suburban, New York high school. Before this policy, the school used a tracking system based on grade point averages and teacher recommendation in order for students to be eligible for an Advanced Placement class; after, students were allowed to enroll themselves without completing any predetermined requirements. This study used ANOVAs and time series analyses to analyze the differences in means of student enrollment and achievement in AP classes among students before, during, and after the policy implementation for all students, as well as for demographic subgroups. Overall, there was a significant increase in the percentage of students who enrolled in at least one AP course for the students who were exposed to the policy and also a significant increase in the average number of AP courses students took. However, when broken down by subgroup, Black and Hispanic students did not realize these benefits. For most classes, AP scores were not influenced by the policy even though enrollment increased. Therefore, the Challenge by Choice policy achieved the school's intended goal of increasing AP course taking and achievement, but also increased equity gaps among some groups. Researchers, school

administrators, and other school personnel may need to consider additional supports to ensure that this policy is effective for all students.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

While the 1950s marked significant gains toward educational equity with the Brown v. Board of Education decision to desegregate schools and the renewed focus on achievement disparities among racial subgroups, the creation of the Advanced Placement system (AP) changed the nature of equity in access to rigorous coursework. Motivated by the desire to challenge the most promising students with college-level curricula, the College Board created and implemented the AP Program nationally. The new AP curricula allowed the "top" students around the country to study college-level material and to earn college credit while still in high school, but also created new divisions along racial and economic lines both between and within schools in who was eligible to take on this challenge.

These divisions are evident in the stratification of access to and participation in AP programs by race, SES, and geography. In 2013, the Education Trust looked closely at AP and IB participation rates nationally and by school. The report found that most schools offered AP programs yet there were still some who did not, resulting in millions of students who did not have access to AP courses in their schools. Similarly, Theokas and Saaris (2013) found that 71 percent of all high schools serving 91 percent of all U.S. students had AP programs. In other words, 1 in 10 students attend a school without AP courses (Theokas & Saaris, 2013). The schools without AP programs tend to be small, high poverty, and rural schools. According to the College Board and Common Core Data from 2010, 74 percent of urban schools and 86 percent of suburban schools have an AP program compared to 59 percent of rural schools. Also, 44 percent of small schools offer

a program compared to 99 percent of large schools (Theokas & Saaris, 2013). Breaking access rates out by students' race/ethnicity and economic status: Asian students (97 percent) were slightly more likely than the average student to have access to AP courses in their high school while American Indian students had the least access (76 percent). Students eligible for free and reduced lunch (91 percent) and middle- and high-income students (93 percent) were similarly likely to attend schools with AP programs and the same was true for Hispanic (93 percent), Black (89 percent), and White (91 percent) students. However, not all schools offer the same amount of AP courses and the racial gaps in access are starker taking this into account. Low-income students are almost half as likely as other students to attend a school with a full course load of Advanced Placement classes. Native American and Black students are also far more likely to have limited course access. Therefore, while many low-income and students of color attend schools that have AP courses, they do not have access to as many AP courses in their schools when compared to the White and Asian students from a higher socioeconomic status (Theokas & Saaris, 2013).

There have been some pushes to reduce these inequities. In 1999, there were two college lawsuits concerning AP classes which argued that minority and low-income students were denied equal consideration in the college admission process because some high schools did not offer AP courses and, in some cases, students could not afford to take the exams. Therefore, the U.S. Department of Education provided $3 million in grants in 1998 to pay for exam funds for economically disadvantaged students taking AP classes (Curry, MacDonald, & Morgan, 1999) and increased this amount to $15 million in the year 2000 ("Promotion of Advanced Placement," 2000). The lawsuits also led to

the California legislature appropriating $30 million to allow access to AP courses online and provide finances to disadvantaged schools to expand their AP programs (Burdman, 2000). Some other states require public high schools to offer a certain number of AP courses or pay for the exams for students (Hebel, 1999). In New York, the Rewarding Achievement (REACH) program was created in 2007 to improve college readiness and completion rates among low-income students in New York City. Students received cash rewards ranging from $500 to $1,000 for scoring a three or higher on any AP exam. Also, money referred to as REACH bonus grants were provided to schools based on students' performance on AP exams. Through federal funding, New York was able to subsidize $1,605, 920 in exam fees for low-income students (McBride Davis, Slate, Moore, & Barnes, 2015),

However, it must be noted that simply expanding the number of schools that offer AP courses or providing supplemental funding does not ensure that students in those schools will have access to or enroll in those courses. Participation in Advanced Placement classes still primarily requires teacher referral (58.8%), participation in pre-requisite courses (53.3%), and earning adequate course grades (49%) or National Exam scores (7%) according to College Board 2008 data (College Board, 2008, p. 19). In other words, in many schools, teachers and counselors serve as gatekeepers to higher level classes.

While this referral system may not be inherently bad, biases and unfair restrictions could limit some students' enrollment in the AP program. Low-income students enroll in AP classes at a rate less than a third of their middle- and high-income peers when they attend schools that offer AP programs (Theokas & Saaris, 2013). If low-

income students participated in AP at the same rate as other students, more than half a million more students would be in those courses. Similarly, the enrollment rates for minority students in AP classes, excluding foreign languages, are low (Whiting & Ford, 2009). Black and American Indian students participated in APs at about half the rate of the national average (Theokas & Saaris, 2013). For example, African Americans made up 9.2% of exam takers but they were also 14.5% of the nationwide graduating class in 2014 (College Board, 2014). They were underrepresented in every AP subject (Kolluri, 2018). If there was equal enrollment, more than 79,000 black students, 37,000 Hispanic students, and nearly 6,000 American Indian students would benefit from access to higher level course works (Theokas & Saaris, 2013).

New York State faces the same equity gaps as the nation in terms of Advanced Placement. In 2018-19, the Education Trust-New York found glaring disparities in both access and enrollment for minorities and low-income students in AP/IB programs. In New York, 85% of schools offer AP/IB courses. However, high-needs school districts are less likely to offer AP/IB programs than low-and average- need districts. Low-need districts offer on average sixteen AP/IB courses whereas rural, high-need districts offer on average four AP/IB courses. Moreover, 98% of low-need and 90% of average-need schools offer AP/IB programs whereas 62% of rural, high-needs schools and 76% of urban, suburban high-needs school offer AP/IB courses. More Black and Low-income students are enrolled in schools without AP/IB programs than any other subgroup (not low-income, American Indian, Asian, multiracial, Latinx, White) ("Course Access: A Statewide Challenge," n.d.).

In schools with AP/IB programs in New York, minority groups are also underrepresented in terms of enrollment in these courses. Statewide, 41% percent of Black and Latinx students go to schools that offer AP/IB courses but only 28% are enrolled. Similarly, 52% of low-income students statewide have access to AP/IB courses but only 36% are enrolled. New York State enrolls students in AP/IB who are not low-income at two times the rate of their low-income peers and enrolls White students at 1.7 times the rate of their Black and Latinx peers. If representation were fair, 198,000 additional students who are low-income and 173,000 additional students who are Black and Latinx should be enrolled statewide ("Course Access: A Statewide Challenge," n.d.).

Teachers, administrators, and other stakeholders must use school policies to address the opportunity gaps plaguing our school systems, especially as the need for a college education increases. A College Board report (2008) urged policy makers in schools to become more proactive and aggressive when addressing AP underrepresentation of certain minority groups within schools (Whiting & Ford, 2009). There are many different types of policies within schools that can address both between and within school opportunity gaps. For example, states must continue to create policy that ensures access to AP programs in places that have "limited resources, geographic segregation, or a lack of will" that limits students' access to the higher-level AP courses (Theokas & Saaris, 2013, p. 9). They should also provide grant programs that give funding for AP/IB programs that focus on equity and equitable participation/enrollment in addition to expanding the number of AP/IB programs. These policy decisions help to increase the number of students who participate in AP and IB programs (Theokas & Saaris, 2013). However, the within school participation gaps must

be addressed also. Policymakers must require high schools to offer a minimum number of AP or IB courses and provide support so these programs can be successful.

Purpose and Study Design

The purpose of this study was to understand the impact of an open enrollment program for Advanced Placement classes, called "Challenge by Choice," on students' enrollment in Advanced Placement courses and achievement on Advanced Placement exams. Specifically, this study examined how particular subgroups within a New York high school were affected by Challenge by Choice. Before Challenge by Choice was implemented, the school employed a tracking policy in which students could only enroll in Advanced Placement classes if they acquired teacher recommendation based on ability, motivation, and interest and a certain grade point average. For students in question or on the borderline, a committee met to discuss the appropriate scheduling for that student. After the policy, students were able to select and enroll in any AP course of their choosing.

I studied the impact of this policy change using ANOVAs and an interrupted time series analysis using secondary AP course taking data from before the Challenge by Choice Policy started and for eight years after the policy was implemented. The primary analysis examined the impact of the policy on the whole student body taking AP classes. However, I also examined certain demographic groups (White, Black, Latinx, Asian, female, male, ELL, non-ELL, students with IEPs, and students without IEPs) during these years to see how the change in policy impacted different groups of students in terms of enrollment and achievement in AP classes.

Theoretical Framework

This study draws on the theoretical frameworks of Marxism and Social Reconstructionism.

According to Bowles and Gintis (1976), The organization of schools function to reproduce inequalities that exist in society. For example, hierarchies of classes in society, similar to regular and AP with prerequisites in schools, function to advance some students and not others. The students who benefit are typically those who have had more opportunities outside of school, such that schools expand existing opportunity gaps. Inequalities in society are also reproduced that those in control resist breaking down (Bowles & Gintis, 1976).

However, Counts (1932) and Dewey (1897) suggest that schools do not have to reinforce societal inequalities and can function to break them down. Counts suggests that school reforms focusing on equity, community, cooperation, and democracy, could help counter societal inequalities between students form different economic backgrounds. Dewey further posited that schools should be the foundation of social change by restructuring society and giving opportunities to all to revitalize democracy. Schools must improve the lives of students through freedom and creativity. Dewey believed that reforms from enactments of law or threats of penalties are ineffective and that the only effective method toward social progress or reform stems from education. Thus, what one wise parent wants for his/her own child should be the same as what the whole community wants for all the children; if this is not the case, it threatens democracy. In schools, individualism and socialism agree. Only through the full growth of all individuals who make up the school can society as a whole advance. Historically, the literate and privileged had the power of possession of knowledge but knowledge now moves through

society in all ways. As such, schools must connect the curriculum to all children and become places that value authentic learning, the interests of students, the importance of communication, inquiry, and self-expression (Dewey, 1899). Through the lens of social reconstructionism, tracking again reproduces the social hierarchy, segregates students, and stifles opportunities to learn. However, opening access and increasing enrollment in challenging classes dismantles the separation and gives opportunities to all students.

Conceptual Framework

There are many arguments for and against a (de-)tracking policy and how it may impact the students. Conceptually, these types of policy decisions can be viewed through Utilitarian and Theory of Justice perspectives. From a utilitarian angle, the value of tracking would be determined by the summing of the benefits of the policy over all students and if the benefits outweigh the costs (net positive), the policy is deemed good. However, this perspective takes no note of the equitability of the policy benefits among subgroups. John Rawls (1971) proposed the Theory of Justice contesting this type of utilitarian consequentialist view. According to Rawls (1971), those who benefit the least from the decision should be given more focus than those who benefit the most. Therefore, from a justice or equity perspective, the relative achievement of students in various demographic groups must be given precedent when comparing a tracked environment to a Challenge by Choice program. We must not consider only the net benefit, but the distribution of those benefits or costs among groups. In this study, I evaluated the implementation Challenge by Choice through a justice lens, identifying not only net benefits but also impacts on individual subgroups to understand its implications for equity.

Significance of the Study

Because of Black, Latinx, and poor students' statistical underrepresentation in AP courses, they are receiving a message that they are not smart enough or that the courses were not intended for them to participate (Patrick, Socol, & Morgan, 2020). By examining one school using ANOVA tests and a quantitative time series analysis, I better understand the benefits and drawbacks of a Challenge by Choice or Open Enrollment policy for encouraging these students to take AP courses. This research helps us understand if this is a model that we can use to reduce opportunity gaps. The implications of this research are far-reaching. It can help administrators, teachers, and counselors plan interventions to change attitudes toward barring certain students out of higher levels classes and school officials can assist with policies that increase enrollment and improve academic achievement on Advanced Placement exams. For example, if there is unequal take-up, counselors can encourage students to enroll in Advanced Placement classes. With a better understanding of differential success once enrolled, teachers can better adapt their classrooms for students of all backgrounds and levels. In the end, students will have more classes that receive college credit, exposure to college level learning, and AP courses on their transcripts which could positively impact their chances for admission into college.

Connection with Social Justice and/or Vincentian Mission in Education

The research in this study directly relates to the Mission of St. John's University in that it examined an issue related to social justice and discrimination. The tracking system has traditionally served as a barrier to some children in the lower ability tracks in receiving the same opportunities as students in the higher ability tracks. Detracking

through policies like Challenge by Choice serves to challenge this institutional structure that bars some students out of challenging courses. Detracking breaks down the technical structure but also the political structures in schools. It confronts issues of power, control, and ideologies of perceived abilities and merit. Detracking efforts often fail because some feel it is in the best interest of students to separate based on perceived cognitive ability; in reality, these ideas are often a façade for wanting to keep students segregated by class and race (Oakes, Jones, Datnow, & Wells, 1997).

Therefore, evaluating a policy such as Challenge by Choice connects directly with the Vincentian Mission of helping those disadvantaged and often discriminated against in tracking policies. Detracking efforts serve to break down these barriers that often keep minority and poorer students from taking challenging courses, such as Advanced Placement. Through these policy changes, the educational system will be improved, particularly for underrepresented and disadvantaged groups.

Research Questions

This study was guided by two questions:

Question 1. How did the introduction of AP Challenge by Choice change student's AP course-taking behavior?

A. How did the introduction of the AP Challenge by Choice policy change the percent of students taking at least one AP course during high school, overall or by student subgroup (White, minority, female, male, ELL, and students with IEPs)?
B. How did the introduction of the AP Challenge by Choice policy change the average number of APs students elected to take during high school,

overall or by student subgroup (White, minority, female, male, ELL, and students with IEPs)?

Question 2. How did the introduction of Challenge by Choice change the trend in AP enrollment and test scores, overall and by student subgroup?

Definition of Terms

Advanced Placement Exam Score. The *r*esults of standardized examinations administered in May of every year after students complete an AP course. The grades are given on a 1-5 scale. For most AP exams, the score is a weighted combination of scores from two sections, multiple choice and free response. Some assessments have other scored sections. The final score is on a 5-point scale that offers how qualified the student is to receive college credit or placement. Each college, however, decides what scores will be granted credit or placement. The AP Program and College Board conducts studies in all subjects to correlate performance of AP students with college students in comparable courses. These studies help to determine the cut points that translate students' composite AP scores into an AP score on a scale of 1-5 ("About AP Scores," n.d.).

Advanced Placement Class Enrollment. Students who enroll in AP courses and enroll in the course via College Board. Students enroll in a particular class section and complete the AP exam at the end of the course ("Student Enrollment," 2021).

Achievement Gap. Achievement gaps exist when one group of students (grouped by race/ethnicity, gender) outperforms another and the difference between the groups' average scores are statistically significant. The academic performance between groups is usually measured on state tests in reading, mathematics, or other subjects, the NAEP assessment, and other norm-referenced assessments such as the NWEA. Academic

performance can also be measured by grades, grade point averages, or other measures (Anderson, Medrich, & Fowler, 2007).

Latinx. In this study, the term Latinx will be used interchangeably with the term Hispanic.

Opportunity Gap. Opportunity gaps exist when different demographic groups experience unequal resources and opportunities educationally causing an uneven playing field for all students to reach success (The Glossary of Education Reform, 2013). For example, Black, Hispanic, and poorer students receive less access to AP courses and challenging curricula compared to White, more affluent students (Barnard-Brak, McGaha-Garnett, & Burley, 2011).

Open Enrollment. A school or district policy that allows students to enroll in any course without prerequisite requirements. Students are not required to meet additional criteria to enter the class. Students only need the motivation and desire to take the class. Many schools or districts are implementing policies such as these to level the playing field and decrease the opportunity and achievement gaps ("More Schools Offering Open Enrollment for AP courses," 2013).

Challenge by Choice. A form of open enrollment policy in which students in this study did not need any prerequisite requirements to get into any Advanced Placement classes. However, for some math classes, prerequisite course completion was required.

Tracking. This refers to a method used to group students according to their perceived ability, IQ, or achievement levels. Students are placed in ability leveled classes to teach them according to their needs. Tracking began in the 1930s but has been controversial in education, especially in the last twenty years ("Tracking and Ability

Grouping in Middle Level and High Schools," 2020). Tracking has been critiqued for providing inequitable opportunities and education to students in lower ability tracks and for separating students based on race and class (Rubin, 2006).

Detracking. A reform in response to the negative impacts of tracking by creating heterogeneous classes of mixed ability students. By placing students in heterogeneous classes, detracking attempts to close opportunity and achievement gaps. Consequently, there are usually political power struggles when trying to implement these reforms because it confronts traditional conceptions of students' abilities and the segregated tracked structure of schools (Oakes, Jones, Datnow & Wells, 1997).

ELL Learner. English Language Leaners, a highly heterogeneous group of students with diverse needs, backgrounds, languages, and goals. In New York State, all newly enrolled students are required to complete a Home Language Questionnaire. If the survey indicates a language other than English is spoken at home, that child must take an English Language Proficiency Test called the New York State Identification Test for English Language Learners (NYSITELL). The results assess the child's English Language level (Entering, Emerging, Transitioning, Expanding, and Commanding). If the child is identified as Entering, Emerging, Transitioning, or Expanding, that child is an ELL. ELLs are entitled to receive Bilingual Education or English as a New Language (ENL) services. The results also guide schools in determining how many minutes of ENL services that child receives ("A Guide for Parents of English Language Learners In New York State," n.d.).

CHAPTER 2

Chapter 1 highlighted the equity concerns within the Advanced Placement program: There is a conflict between the lack of access of AP courses for some students and the growing need to take such classes to enter college or university. The onus is on schools to develop a system and policy to ensure all students have fair opportunities to learn, which starts with equal access to advanced programming. In this chapter, I lay out the theoretical framework for my study, provide a brief history of Advanced Placement program, and review several related research studies that describe the positive and negative results of both tracking and detracking policies.

Theoretical Framework

Marxist, social reconstructionist, and constructivist theories provide a frame for understanding policies related to offering Advanced Placement classes to all students.

Marxism and Education

Bowles and Gintis began the publication of *Schooling in Capitalist America* in 1968 amid a growing academic debate about the purpose and structure of education. Their positionality focused on a sphere of education that would be free of prejudice and full of equal opportunities and materials. The failing of U.S. education sparked their initial work and continuation afterward. Bowles and Gintis argued that although cognitive skills are important in the economy and predicting economic success, cognitive development is only partially developed in schools. Instead, schools focus on reproducing a human workforce that fits within a hierarchy of modern corporations. For example, the workforce must learn to continually do work without complaint and underneath a command of people. They argued that schools engrain this hierarchical structure in

students at a young age teaching compliance instead of questioning and completion instead of creativity (Bowles & Gintis, 1976).

Bowles and Gintis (1976) also explained that economic status is often passed onto children despite the studies before their book that claimed the opposite. When the errors from the prior studies were corrected, the correlations became quite substantial. The reason for parents' economic status being passed down to children is due in part to the unequal educational opportunities and superior education services provided to wealthier children and the economic advantages offered to children from higher socioeconomic statuses. Those students with more schooling and resources had higher earnings in their future, thus reproducing social inequalities (Bowles & Gintis, 2001).

Moreover, personality traits, rather than skills, are determinants of success in the workforce. Ultimately, those behaviors are learned and reinforced in school. Schooling does more for students than just enhance cognitive skills. This was referred to as the Correspondence Principle (Bowles & Gintis, 1976). The Correspondence Principle argued that schools produce future workers and socialize students to accept certain beliefs, values, and behavior based on authority and a hierarchical structure. In contrast, students are not encouraged to question, criticize, or focus on their interests. Schools do this in two ways: First, schools influence students' exposure. Teachers, who may not be representative of the population of parents or students, act as behavioral models for the students. For example, students learn certain cultural traits from their parents, at first, and then learn from teachers when they enter school. If the traits are different, from parents and teachers, the child must choose which traits to adopt. To determine which cultural traits to adopt, students consider the rewards and punishments associated with school.

Switching cultural traits increases as the net rewards increase that the students experience. Therefore, schools can promote traits that are beneficial to one group in power even if those traits are not beneficial to that group adopting them (Bowles & Gintis, 1976).

Secondly, schools offer rewards and punishments to students for certain behavior. There is a close association between personality and behavior traits of students with good grades in school and the traits associated with getting high supervisor rankings in the workforce. For example, subservience and compliance is beneficial to the employers or the teachers at the expense of the employees or students. These behaviors stifle critique and creativity but are rewarded to make it easier for the leaders in the hierarchical structure. Therefore, schools have impacted students both cognitively and non-cognitively reinforcing the societal hierarchies for generations (Bowles & Gintis, 1976).

Social Reconstructionism

Social Reconstructionists called on educators to fix the inequities in societies produced by capitalism. Prior to the Industrial Revolution, education was a very selective process only for the top echelon of society (Counts, 1926). In the beginning of the twentieth century, public high school education expanded the curriculum and opportunities offered to all students not just an elite subset. Theoretically, equality of opportunity then became a part of the American education system; students were not tied to the inevitability of their economic classes (Counts, 1926). However, while schools opened their doors to all, the social stratification in society was mirrored in schools and schooling. By not allowing certain students to access top programs, the democratic public school system did not provide equitable opportunities to all students. Moreover, the top

economic and social classes usually went to school for longer periods of time (Counts, 1922).

Counts also wrote about the close relationship between perseverance in school and parents' occupations (Counts, 1922). He posited that higher social classes dominate boards of education and school policy makers. Because schools are often run by people who have social and economic power, practices and policies usually remain aimed at keeping the status quo to preserve those positions of power (Counts, 1928). For example, boards of education are very selective and usually composed of educated people from higher classes. In turn, this economic class controls the policies of the school and thus the economic resources of the country (Counts, 1928).

Consequently, Counts (1927) argued that schools indoctrinate students to maintain the social hierarchies. He found that lower class families enrolled in more practical courses teaching skills that students could use right after high school. In contrast, students from upper levels of society chose courses that prepared them for college indicating the selectivity of curricula in high schools. Accessibility was key for students however access was provided by those in control. He said, "only the most stupid and unenterprising can fail to perceive the promise of power which the school holds out to those who would organize its curriculum; only the most public spirited can refuse to yield to the temptation to use such power in the attainment of selfish ends..." (Counts, 1927, p. 82). Those in power-controlled curriculum and policies for their own selfish means and to reinforce their positions of power.

Despite the bleak assessment, Counts also asserted that teachers and schools had the power to help the masses and reinvigorate democracy. Rather than a focus on

individual success, competition, and property rights, education would focus on social reform, the community, and cooperation. Through these ideas, students could take control by becoming dedicated to cooperation rather than exploitation. This focus in the curriculum could reform society and form a new social order. Progressive education, he argued, had to "emancipate itself from the influence of class" (Counts, 1932, p. 259). In order to do this, education had to rethink its purpose; teachers had to take control of the curriculum, meld the issues in society into the curriculum, and provide access and equality of opportunity to students regardless of their socioeconomic status. Through these changes, leadership would be attained by merit, not by background (Counts, 1932).

Educational philosopher John Dewey also described schools as agents for change. He believed that when schools level the playing field for disadvantaged students so that they have access to and are enrolled in the best curriculum, they too have possibilities in their future. The purpose of schools, according to Dewey, is not to retain the status quo of social stratification so that only the top would succeed. Dewey believed that school is a social institution, and that education is a social process (Dewey, 1897). As an active, constructive process, effective teaching and learning promotes critical thinking instead of the acceptance of content knowledge. Unfortunately, the traditional classroom setting does not allow for the natural, questioning spirits of developing children. Therefore, the classroom helps the system and not the individuals inside it. Moreover, the trouble lies in that each generation tries to maintain existing conditions and keep the status quo. But through education, society could shape itself and move in the direction it wished to move. As an institution, Dewey posited that schools are the most effective and primary forms of social progress and reform (Dewey, 1897). Dewey contended, "Parents educate their

children so that they may get on; princes educate their subjects as instruments of their own purposes" (Dewey, 1958, p. 111). Only through educational reform could society create a true democracy, however without constructive reflection of the system, education could also merely perpetuate the societal problems that already exist (Dewey, 1958).

Unfortunately, traditional schooling has continued the indoctrination of the current social hierarchy. Parents, especially from the upper classes, are often the defenders of the status quo within schooling and that the accepted social and moral values be passed down to children. In order to reach a true education and democracy, schools must have a new, unified objective and give the oppressed or disadvantaged groups an equal share of control. The first way of doing this, according to Dewey, is to root all education in current social problems. Through this type of inquiry learning, the individual learners become aware of their roles in society and how their actions either perpetuate or change the current social order. According to Dewey, participation is also a big part in achieving social change because only those who participate together toward common goals can help achieve true democracy and create more engaged citizens. Equal participation of all students in this type of learning is paramount to personal and societal growth. Education, therefore, increases the capacities of individuals but this in turn should benefit the community and society at large. Personal improvement should be applied to the improvement of society for all (Dewey, 1937).

Within the context of today's world, the struggle to democratize education still exists. Restricting access to certain curricula to some students is still an issue of paramount importance as is recreating the societal social strata within schools. Whereas the education system in the early twentieth century was just being established and now all

must have access to schooling, the inequities between certain groups within schools are still there. How we achieve equity for all students must start by looking at the policies put in place to either promote or inhibit it.

Conceptual Framework

Guided by the theoretical foundation that schools have the potential to either reproduce or reduce systemic inequality, we can evaluate existing and new policies that govern students' experiences in schools. In the context of this study, I focused on tracking and detracking policies. The effect of any given policy can be evaluated along two dimensions: First, we can evaluate a policy's utility—its ability to produce *net positive outcomes* for students. Second, we can examine its equitability or ability to *reduce systemic inequalities* evident in the education system.

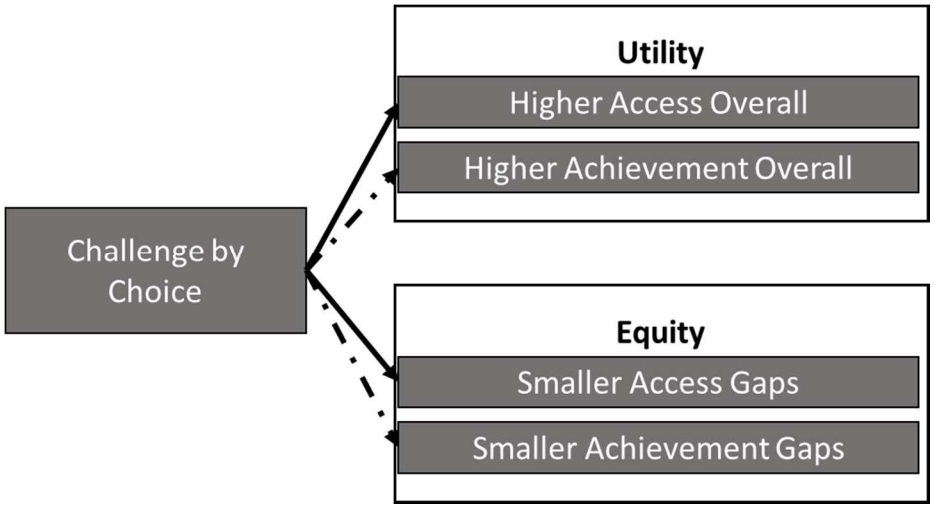
According to the utilitarian philosopher John Stuart Mill, each action produces consequences that make an action right or wrong. Actions considered right produce positive net outcomes for the greatest number of people. Each person is counted exactly as much as another. Guided by this principle, utilitarianism looks at what helps society as a whole and not necessarily which groups receive the most benefits (Mill, 1863). Consequently, in the context of schooling, a policy, such as Challenge by Choice, could be deemed beneficial (according to utilitarianism) if it produced net positive outcomes for all students but benefitted different subgroups unequally.

Herein lies the challenge of looking at utility alone. The world is not equitable, and existing inequalities in opportunities can easily be overlooked or exacerbated by focusing only on net gains. To contest utilitarianism's lack of equity, John Rawls (1971) created the Theory of Justice. Rawls argued that institutions should focus on benefitting

the least advantaged members of society, the Maximin Principle (Rawls, 1971). In effect, this principle requires comparing levels of utility between people and focusing on those whom the lowest level(s) of utility values. As such, the economy could work for everyone's benefit and those born with natural potential do not continue to get richer at the expense of those who are less fortunate (Rawls, 1971). In the context of a school policy, this would argue for focusing on how any given policy affects the relative standing of subgroups. A policy would not be fair or just if it exacerbated existing inequalities among groups, for example in access to rigorous courses or in test scores.

The conceptual model for the present study is shown in Figure 1, below. I examined the effect of the implementation of a Challenge by Choice policy (as an alternative to tracking) in terms of utility and equity. I conceptualize that there are two dimensions that such a policy would affect: access to rigorous coursework and subsequent achievement on the AP tests.

Figure 1
Conceptual Model



Moving from a tracked environment to a Challenge by Choice environment has clear implications for students' access to rigorous coursework, thus represented by solid lines in the framework. Tracking restricts access to certain Advanced Placement programs for students and segregates students based on race, color, and SES. Challenge by Choice allows students from all backgrounds the opportunity to take rigorous courses; however, the policy falls short of requiring these courses or ensuring that students of different backgrounds feel prepared to take them. This study evaluates whether the offer of an advanced course was sufficient to shift inequities in access to rigorous courses and who was still left behind.

It is not clear how a Challenge by Choice policy could affect achievement and achievement gaps, thus represented by dotted lines in the conceptual framework. First, it is important to note that achievement is conditional on enrolling in the AP courses. Achievement might have decreased because more "low achieving" students (students who previously would not have met academic standards to enroll) were in the classes. Because prior achievement is highly correlated with future achievement, AP scores could have dropped post-policy. However, even if the average scores dropped, more students might be earning AP credit that would count in college.

Second, students' experiences in the classes themselves also affect their outcomes. The narrative that has supported tracking purports that these students would not be able to succeed in more rigorous classes. However, learning theories suggest that may not be the case. According to Vygotsky, learning creates the zone of proximal development, or the information people can internalize with the help of those around them. Only through the help of peers and others in the environment can anyone begin to

learn new information (Vygotsky & Cole, 1981). Any learning a child or adult does comes from previous experience. Children learn by speaking to adults and by imitating adults; therefore, learning and development are always connected and interrelated. According to Vygotsky, students have an actual developmental level that can show their mental development. However, more importantly, what the child can do with the assistance of others might be more indicative of their abilities. The zone of proximal development is the difference between the actual development level determined through independent problem solving and potential development if assisted by peers or adults. This theory asserts that good learning and instruction advances development (Vygotsky & Cole, 1981). Consequently, learning results from experience, engaging peers, and building on previous levels of knowledge. Only in classes that provide students to reach this zone of proximal development can true learning occur. By building on previous schema, students can develop further especially if they are helped by those around them in challenging classes. Also, students learn from others in the school environment; this relationship helps them to construct new knowledge based on experiences and conversations with those around them, an idea to consider when creating classes that only accept certain students into them. According to these ideas, learning is a social activity impacted by students' environments. Therefore, barring students out of higher-level classes leaves them bereft of certain interactions and learning environments they could grow from. Policies such as tracking result in the loss of opportunities for students to learn from each other socially. To fully understand the consequences of the Challenge by Choice Policy on achievement, I looked at the absolute gains in achievement scores in

my study. However, I did not have the power to determine if achievement gaps between subgroups would be improved by the Challenge by Choice policy.

History of Advanced Placement

The Origins of Advanced Placement Classes

The modern Advanced Placement (AP) Program was born out of an effort to maintain competitiveness during the Cold War race to produce the best, brightest students in the world. The program brought together elite high schools and colleges in order to challenge the top high school students by providing them college level curricula. In 1950, John Kemper, the Headmaster at Philips Andover, a private, preparatory school, worked with two other private secondary schools, Exeter and Lawrenceville, in conjunction with Harvard, Princeton, and Yale Universities with funding from the Ford Foundation's Fund for the Advancement of Education (FAE) to increase academic rigor. The FAE in 1951, primarily focused on gifted and talented student education so this project suited the FAE's goals and mission. This project soon became the School and College Study of Admission with Advanced Standing, which was led by twelve colleges and universities and thirteen secondary schools. The study's 1952 report suggested that the U.S. challenge the most motivated students in chosen, prestigious secondary schools so that those students could earn college degrees in less time, allowing those able-bodied students to enter the workplace or graduate school sooner. Over time the AP program became a mark of prestige for the students who took those classes (Schneider, 2009). In 1954, 532 students from 18 secondary schools took 929 placement exams; in 1955, the Advanced Placement program was established under the leadership of Charles Keller and in 1956,

the first exams were administered under the organization, the College Board (Schneider, 2009).

Therefore, the intent of the Advanced Placement program, initially, was to challenge only highest achieving students and enable them to earn college credit. Even the second director of the AP program, Dudley, conveyed that students were not equal and not all should have access to the program. In the following years, the program grew very rapidly. It was no longer restricted to only students at Ivy League schools, and it became a key factor in acceptance to the top colleges or universities (Schneider, 2009). In 1961, educator and author, James Conant, wrote in his book *Slums and Suburbs* that the top schools who admitted students who took AP courses in high school were MIT, Michigan, Stanford, Northwestern and six Ivy League schools. At Harvard, about half of the class had participated in the program. In other words, for many students, taking AP classes became similar to attending a preparatory school (Conant, 1961).

The Growth of the AP Program and Equity of Access

During the democratic movement of the 1960s and particularly Lyndon Johnson's Great Society reforms, educators began to shift their focus toward students who were often denied a challenging high school education and opportunities to attend post-secondary education (Schneider, 2009). The fact that Advanced Placement helped the top students but also created vast inequality became more obvious as AP classes became more necessary for college admissions. The egalitarian reforms of the 1960s highlighted that the opportunities Advanced Placement offered for the top students to get into college were only available for those that had access to it, primarily white students from the suburbs or from private schools (Schneider, 2009). However, the reform effort took a

setback when “A Nation at Risk” was published in 1983 and challenging the most elite students was again at the forefront of the educational agenda. It was not until the 1990s when the education ideology would again reverse and federal funding would be used to help more communities provide AP classes to their students (Klopfenstein, 2004a as cited in Kolluri, 2018).

The expansion of the Advanced Placement Program continued as the value of the college degree continued to rise (Kolluri, 2018). The Lumina Foundation in 2012, predicted that by 2025, 60% of new jobs would require a post-secondary education degree. A 2013 study of 1,380 colleges and universities by the College Board found that 68% of their school policies offer credit for a score of 3 or higher on an AP exam, 30% offer credit for a score of 4 or higher, and 2%, only offer credit for a score of 5. Eight (<1%) of institutions did not accept AP course credit (Adams, 2014 as cited in Kolluri, 2018). In 1954, 523 students took at least one AP test and fifty years later, that number reached 1.1 million. By 1976, 3,937 schools and 75,651 students took part in the AP program and by 1985, 6,720 schools and 205,650 students participated in Advanced Placement classes. By 1994, 11,500 schools and 458,945 students participated in the AP program (Schneider, 2009). In 2016, more than 2.5 million high school students took more than 4.7 million AP tests, up from less than 4.5 million the previous year (College Board, 2016a).

In terms of equity, minority student enrollment in Advanced Placement has risen as well. Minority students comprised 19.5% of students taking AP tests in 1988, twice the number since 1983. By 1994, 26.3% of AP Test takers were minority students. Some states encouraged AP growth by providing grants to increase and support the programs.

The U.S. Federal government, in 1998-99, spent $2.7 million in subsidizing exam fees for students and professional development for teachers who needed it. In 1994, 14.9% of students graduating high school had AP credit but by 2013, that number increased to more than 39%. African Americans have increased their participation rate by 200% from 1994 to 2013 to a participation rate of 27%. Thirty-six percent of Latinx students in 2013 graduated with an AP course credit compared to 41% of White students (Malkus, 2016 as cited in Kolluri, 2018). Access to AP for students whose parents did not graduate from high school also has expanded from 8.4% participation in 1994 to 29.1% in 2013 (Kolluri, 2018).

Despite this growth, there are still documented inequities. Zarate and Pachon (2006) used data from the California Department of Education's California Basic Educational Data System (CBEDS) to analyze AP course offerings among 1,094 (n = 1,094) public high schools. They showed that from 1997 to 2003, the average number of AP courses offered per school increased from five to six courses; however, the percentage of schools not offering AP courses rose from 15 percent in 1997 to 22 percent in 2003. They further found that the number of African American, Latinx, and low-income students was negatively correlated with the number of AP course offerings in a school. For example, on average, large schools with high concentration of minorities (75-100 percent) offered nine AP courses, while large schools with low concentration of minorities (10 percent or less) offered 13 courses, a 28% difference. Among medium-sized schools, schools with less than 10% of African American and Latinx students offered more than 1.5 times as many AP courses compared with schools that served 75% to 100% African American and Latinx students (Zarate & Pachon, 2006).

Similarly, Barnard-Brak, McGaha-Garnett, and Burley (2011) showed the inequity in access to Advanced Placement courses for certain students in the United States, as a whole, by examining a sample of 12,144 (n = 12,144) children from the NELS (National Center for Education Statistics) in the year 2000. The authors concluded that school characteristics, such as the percentage of minority students and the percentage of students who received free and reduced lunch, were inversely associated with the number of AP courses offered at their schools. These characteristics were not significantly associated with the number of students enrolled in the AP courses however (Barnard-Brak et al., 2011).

While the College Board's current objective is to provide access to a high-level college curriculum to all students, regardless of background, schools' course assignment or tracking policies may not make this possible. Schools must, therefore, self-examine to understand if their policies are inhibiting equity in access to high-level coursework.

Benefits of the Advanced Placement Program

Importantly, research studies have delved into the overall academic benefits of the Advanced Placement program for students and thus the importance of closing opportunity gaps. One study by Ackerman, Kanfer and Calderwood (2013) found that there was a positive impact on college exam scores for students who were enrolled in Advanced Placement courses, college GPAs, and college graduation rates. The sample consisted of 26, 693 students (n=26,693) who entered the Georgia Institute of Technology as first year undergraduate students from 1999-2009. The results of the study found that increased numbers of AP based course credits were associated with higher GPAs at Georgia Tech for the first year and beyond. Students with AP credits tended to

complete a greater number of higher-level courses. Those students also graduated at a higher rate and graduated in fewer semesters. Average AP exam score was the best predictor of academic success after students' high school GPAs. Additionally, this study showed that the important predictors of STEM major persistence were receiving credit for AP Calculus and if the student had completed three or more AP exams in the STEM area successfully (Ackerman, Kanfer, & Calderwood, 2013).

Flowers (2008) also found a positive impact on college entrance exam scores, college GPAs, and rates of obtaining advanced degrees for those students enrolled in Advanced Placement programs. Using data from the National Education Longitudinal Study (1988) and then subsequent surveys in 1990, 1992, 1994, and 2000, Flowers (2008) found that the respondents who participated in AP programs were more likely to score higher on college entrance examinations than those who did not participate in AP programs. African American and Hispanic students who participated in AP programs scored 100 points higher on college entrance exams than African Americans and Hispanic students who did not participate in AP programs. The results also showed that respondents who participated in AP programs reported higher undergraduate GPAs than students who did not participate in AP programs, particularly for Asian or Pacific Islander and White students. Furthermore, Asian or Pacific Islander and White students who did not take AP classes were less likely to earn an advanced degree than African Americans who did not participate in AP. Hispanic and White students who participated in AP were more likely to earn a bachelor's degree and an advanced degree than Hispanic and White students who did not participate in AP. There were also income differences according to AP participation. Specifically, Asian or Pacific Islanders and Whites who

participated in AP programs reported earning about $10,000 and $3,000 more a year respectively than the same groups who did not participate in AP programs (Flowers, 2008).

Moreover, Hansen, Reeve, Gonzalez, Sudweeks, Hatch, Esplin, and Bradshaw (2006) found a positive impact for those students who took Advanced Placement courses on writing ability in college. The study looked at two three-page writing samples from 182 sophomore students enrolled in a history of civilization course. The essays were scored by two readers; the results showed that the students who had taken an AP English course and a First-Year composition course performed significantly higher than those who had only taken AP English or First-Year Composition alone. While the study shows that exempting students from first year writing courses in college may not be the best decision for students, it also indicates that AP courses did prepare students for college writing. The writing and college preparatory work may also help students in other college classes (Hansen, Reeve, Gonzalez, Sudweeks, Hatch, Esplin, & Bradshaw, 2006).

Lastly, Geiser and Santelices (2004) found that students who took AP courses had an increased chance of being accepted into college and that AP exam scores strongly predicted success in college. The study examined Advanced Placement criterion for admission into the University of California and found that the number of AP courses had little to no impact on students' performance in college, but AP exam grades were a significant predictor of success in college courses. The study also found that selective colleges and universities gave special consideration to students in admissions, but the degree varied among the college or university. For example, the University of California recalculated students' high school grade point average and gave additional points for AP

or honors courses. Others accepted district weighting of AP courses; the most widespread practice was that admission officers used the number of AP/honors courses as part of a comprehensive overview of a student. Some colleges or universities gave extra points for AP or IB courses while others compared the records of students from the same high school to see what the student took advantage of that was available to them. However, the findings of this study revealed that institutions might need to reconsider AP as criterion in high stakes admissions because of the disparities in access to AP courses for disadvantaged and minority students (Geiser & Santelices, 2004).

Therefore, studies show that some of the academic benefits of taking Advanced Placement courses include a better chance of admission into college, success in college, and obtaining advanced degrees after college (Ackerman et al., 2013; Flowers, 2008; Hansen et al., 2006; Geiser & Santelices, 2004). These profound advantages speak to the importance of schools assessing and addressing the within school opportunity gaps in terms of enrollment and/or achievement of students, particularly those often underrepresented, in Advanced Placement courses.

Review of Related Research

Literature on tracking and detracking students can contribute to this debate by analyzing the effects of such policies on student's access, enrollment, and achievement in courses. This study examines the policy of Challenge by Choice, a form of detracking for Advanced Placement students, and its impact on student enrollment and achievement. Prior to this policy, students had to meet certain criteria to access Advanced Placement classes, however with this change came open enrollment. While the policy at the core of this study may differ in keyways from traditional tracking policies, we can learn from the

impact of opening enrollment to all students by comparing tracking results to open enrollment results.

Why Track?

Academic Outcomes. Overall, studies of the effects of tracking on academic outcomes have shown mixed results. Some studies have shown that tracking can enhance academic achievement, particularly for higher achieving students (Duflo, Dupas, & Kramer, 2011; Epstein & MacIver, 1992, Figlio & Page, 2002; Loveless, 2016). For example, Figlio and Page (2002) found that tracking helped higher achieving students and did not harm lower achievement students. The authors used the National Education Longitudinal Study and analyzed the change from 8th and 10th grade in the student's item response theory (IRT) math score. They then compared students who were taught in a classroom of similarly skilled students (tracked) to those taught in a classroom that had a large variance of student abilities (not tracked). Because of potential endogeneity issues, they compared estimates recovered when using different definitions of tracking (e.g., based on teacher reports vs. principal reports). The results were largely consistent across the different specifications of tracking tested (Figlio & Page, 2002).

Loveless (2016) further found that states that tracked more students in eighth grade math scored better on their AP exams. Using 2009 NAEP data, Loveless found that on average states tracked about 75% of eighth graders in math with some learning algebra and others taking another year of general math; Arkansas had the least amount of tracking (50%) and Nevada the most (97%). Using 2013 AP data, Loveless then showed that states with more tracking had more passing scores. Utah tracked 89% of students in 8th grade math and 70% of AP students scored a three or higher. However, in Texas, only

57% of students were tracked and only 52% of the AP students passed. Even when controlled for socioeconomic status and minority status, Loveless still found a relationship between tracking and AP scores. Moreover, this finding held true for Black and Hispanic student subgroups within states (Loveless, 2016).

However, other studies showed that there may be "winners and losers" when tracking (Argys, Rees, & Brewer, 1996, p. 624). For example, Argys, Rees, and Brewer (1996) used data from the National Education Longitudinal Study of 1988 (NELS), a nationally representative student survey conducted by the National Center for Education Statistics (NCES), to study the impact of tracking on student achievement, controlling for class sizes, teacher experience, and teacher education. Track was determined by teacher report, who stated if their class was composed of students of above average, average or below average or widely different achieving levels relative to other students in the school. The final sample consisted of 3,405 students (n = 3,405) who attended public schools in both eighth and tenth grade and for whom all information was available for the study. Evidence showed that the 8th grade track placement had an effect, regardless of test scores, on tenth grade track. Socioeconomic status was also related to track placement; students with higher SES backgrounds were more likely to be placed in upper-level tracks. In contrast, race and ethnicity were not good predictors of track placement. Students in the lower-level tracks gained a lot from detracking efforts at the expense of the students in the upper-level tracks. The typical student, if placed in a heterogeneous class, would score 63.08 which was more than a percentage point lower than the sample mean of 64.46. However, the effect varied by prior achievement. The estimation results showed that moving from a class composed of below average students to a heterogenous

class led to an increase of 8.6 percent in scores, whereas moving from a class composed of above average students to a heterogenous class led to an 8.4 percent decrease in scores. Thus, continuing to track benefits the highest achievers at the expense of the lowest (Argys et al., 1996).

In a related study, Lavy, Passerman, and Schlosser (2008) studied whether unusually low achieving classmates had any effect on the outcomes of "regular" pupils. The low ability students were determined if their birthdate was before the date for their cohort meaning they repeated the grade. The data used was from administrative records collected by the Israel Ministry of Education for 7 consecutive cohorts (1994-2000) of tenth grade students. The results showed that the proportion of low ability students in class lowered the achievement of regular students. The negative effects were moderate: A one standard deviation increase in the proportion of repeaters led to a decrease of .012 and .036 of a standard deviation in the average score of regular students for high school (Lavy, Passerman, & Schlosser, 2008). This study therefore suggests that regular or high-achieving students benefit from tracking that separates them from the lowest achievers.

Finally, some studies found no effects of tracking on academic outcomes. Using data from Chicago Public Schools, Lefgren (2004) found that the achievement between low and high achieving students was not different in schools that track than in schools that do not track. The sample consisted of 80,003 third grade (n= 80,003) and 94, 230 (n=94,230) sixth grade students. Over 80 percent of the students were either Black or Hispanic and somewhat low achieving. Lefgren constructed a tracking measure based on the initial ability of a student and the ability of classmates since informal tracking often occurs and tracking tends to differ across schools. The study found that many schools do

not track since there was very little overlap of initial test scores, and that tracking had little impact on third grade achievement and sixth grade reading. However, tracking did appear to be associated with increased sixth grade math performance (Lefgren, 2004).

Non-Academic Outcomes. The emotional well-being of students in tracked and non-tracked classrooms has also been debated. Kulik and Kulik (1992) presented the results of a meta-analysis of research on grouping programs falling into five categories: multi-level classes, cross-grade grouping, within-class grouping, enriched gifted and talented courses, and accelerated gifted and talented courses. For multi-level grouping analyses, 56 studies were examined. Only 13 out of 56 studies described effects of ability grouping on self-esteem. The average effect size was 0.19 for lower ability students, -.09 for middle ability students and -.015 for higher ability students. Therefore, instruction in homogeneous classes seemed to raise self-esteem scores of lower ability students and lower the self-esteem of higher ability students. The overall effect size of grouping on student self-esteem in the 13 studies was to decrease self-esteem scores by .03 standard deviations, an insignificant amount. For the enriched classes for the gifted and talented students, five of the twenty-five studies investigated the program's effects on self-concept. The results showed that in four out of five studies, self-concepts were better when the gifted and talented were taught in separate groups. The average effect size was small ($d = 0.10$). Lastly, for accelerated classes for the gifted and talented, the eleven studies showed that on average, acceleration had little or no effect on students' attitude toward school, participation in school activities, popularity, or adjustment (Kulik & Kulik, 1992). Overall, the effects of grouping on students' self-esteem were near zero. They were slightly positive for lower ability students and slightly negative for higher

ability students. Higher ability students may have been less satisfied with themselves around peers similar to them and lower ability students may have slightly gained self-confidence when taught with others at the same ability level (Kulik & Kulik, 1992).

Another study by Wiatrowski, Hansell, Massey, and Wilson (1982) explained that tracking placements had no correlation to delinquency for students. This study used a multivariate model to examine the effects of tracking on delinquency in a large, national longitudinal sample. Five longitudinal sets of interviews were collected from a selection of high school males. Interviews and surveys included tests of ability and academic skills, questions about family characteristics, measures of self-concept, values, educational and occupational plans, and self-reports of delinquent and nondelinquent behaviors. The data was examined to assess the changes in delinquency caused by curriculum tracking from sophomore to senior year and also one year after graduation (Wiatrowski, Hansell, Massey, & Wilson, 1982). Curriculum placement was correlated -.07 with delinquency measure in tenth grade, -.10 with delinquency measured in senior year, and - .09 with the frequency and -.08 with the seriousness of delinquency one year after high school. These correlations were all significant but so small that there was a negligible tendency for college bound students to be less delinquent than non-college tracked students. Therefore, the results showed that tracking had large total associations with many variables, but delinquency was not one of them. Also, the inclusion of curriculum tracking was not significantly predicted by program track when prior levels of delinquency were controlled. Even with initial levels of delinquency excluded, tracking did not significantly predict delinquency in students' senior year or the year after high school (Wiatrowski et al., 1982).

Lastly, a study by Damon Clark (2010) found that tracking impacted whether students went to an elite school. Clark examined the impact of attending a selective high school in the United Kingdom on test scores, high school course taking, and enrollment in a university. Students were assigned to certain schools based on primary school test scores. He found that there was a huge peer advantage of attending a selective school, however four years of attendance in a selective school had small effects on test scores. Selective schools did have positive effects on course taking and university enrollment. If the goal of a public school system is to have positive impacts on what courses students take and the universities students attend, this study suggests that exposing students to higher level courses can lead to these positive impacts on students. These results, however, might not be generalizable to a public-school setting that has a more heterogeneous population (Clark, 2010).

Teachers' Perspectives. Another tracking benefit, according to some teachers, is that heterogenous classes are difficult to teach and students receive a less adequate education compared to classes grouped by ability (Rosenbaum, 2000). Rosenbaum (2000) examined the experiences of teachers in a detracked high school and questioned the practicality of detracking in the classroom. The data for the study came from interviews with eight of the ten social studies teachers from a suburban, high school in the Midwest. The school was relatively homogeneous and students came from mostly White, working- and middle-class families with a few from low-income families. The entire school went from a tracked system with three ability groupings to a completely detracked system, except for foreign language and math which were taught at different levels. The results of the study showed that teachers faced irresolvable conflicts because

detracking pulled them in so many different directions. In the tracked environment, teachers could focus on one group of students. However, in a detracked setting, when a teacher helped one group, the other group of students became disengaged. Therefore, teachers tried to teach to the middle of the class but as they did, they realized they were losing students at both extremes. They also found they had to delete many of the topics covered in upper track classes and use language that was below the vocabulary of higher ability students that they used in upper ability tracked classes. There was a constant struggle over the pace of the lessons as well. If the pace was too fast, slower students would be confused and if it was too slow, the faster students would become disengaged. Thus, in this environment the faster students were again usually the "biggest losers" but in a tracked environment, students' needs could be met (Rosenbaum, 2000).

Why Detrack?

Although tracking might be helpful in a utilitarian sense, its implications have produced inequitable school programs. The gatekeeping tactics barring some students from challenging classes has consequently caused a separation of students based on race, ethnicity, and SES (Burris & Garrity, 2008). To mitigate these inequitable results of tracking, detracked policies have positively impacted students and their social and academic progress. Fully detracked schools have found success in terms of enrollment, achievement, and future decisions of students (Burris, 2008; Alvarez & Mehan, 2006; Dougherty, Mellor & Jian, 2006).

Opportunity and Equity Gaps. The lack of access to certain classes in tracked schools is pervasive both between and within schools in the United States. One study showed that access to certain classes depended on the type of school students were

registered in (Domina, Hanselman, Hwang, & McEachin, 2016). According to Domina, Hanselman, Hwang, and McEachin (2016), high achieving, more affluent schools promoted tracking in more cases while the lower achieving, poorer schools, detracked more. The study analyzed 1,524 (n=1,524) California schools from the years 2003-2013, the period in which the state pursued an Algebra for all initiative for eighth graders. The results of the study show that most California schools created a higher track for students. Socio-economically advantaged, high achieving and academically heterogeneous schools were more likely to take advantage of this option. Many state and educational leaders, at this time, pushed schools to detrack students by enrolling high and low achieving students into 8th grade algebra courses together. Interestingly, however, California schools expanded their mathematics curricula by enrolling a select group of students in advanced geometry courses; this approach was especially prevalent in schools with advantaged populations. Rather than creating heterogeneous learning environments, high achieving schools found a new way to intensify curricula for elite students. Highly socio-economically disadvantaged schools were more likely to pursue Algebra for all strategies and less likely to track up like affluent schools did. The detracking initiative could have eliminated the inequity in access by allowing all students the opportunity to all pursue Algebra, the original intent of the program (Domina et al., 2016).

Another study examined three four-year senior high schools to understand how educators framed tracking decisions. One school served a racially and socioeconomically diverse group of students, another served a middle to upper middle-class White and Asian population, and the third was nearly all composed of African American and Latinx students, many of whom were poor. The results found that schools viewed students'

abilities and motivations as fixed. One principal explained that he could tell by the end of kindergarten which students would be successful in high school. Very few teachers (six at one school, two at another, and seven at the last) recalled any circumstances in which students significantly changed or improved academically so that they could switch tracks (Oakes & Guiton, 1995). As such, tracking efforts closed access to students within the schools themselves. However, the data also showed that the school that was more affluent offered more Advanced Placement and honors courses and fewer low-level courses than the other two schools. Lastly, race, ethnicity, and social class signaled ability and motivation and thus impacted curricular tracks (Oakes & Guiton, 1995, p. 15). The middle class, White and Asian students were seen as high achieving and highly motivated and were therefore offered a rich curriculum. The students in the racially diverse school were seen as diverse in achievement and motivation and thus, offered a college-oriented curriculum with fewer advanced courses than the previous school and many vocational options for average students who probably wouldn't attend college. The administration at the school with mostly poor Latinx and African American students offered the fewest college preparatory classes and had the smallest offering of vocational programs. African American boys at this school were more than twice as likely and girls more than four times as likely to take six or more vocational courses than the racially diverse school (Oakes & Guiton, 1995). Therefore, tracking impacted access to coursework both within and between the schools. Although this data clearly links race, ethnicity, or socioeconomic status to curriculum offerings, only one educator indicated that this might be true. Others cited motivation, student choice, or ability as the reasons for certain curricula placement (Oakes & Guiton, 1995). Detracking efforts would force

educators to change those mindsets and offer the same amount of challenging coursework to all students.

Similarly, another study by Spade, Columba, and Vanfossen (1997) found that course offerings were more constrained in working class schools compared to middle class and affluent schools. Using a series of regression analyses, the study used three pairs of schools from affluent, middle class, and working-class districts. The results found that the higher the social class of the district, the greater the number and diversity of college preparatory or advanced courses. In terms of enrollment, the lowest social class districts and the average (in terms of test scores) schools were less successful in enrolling students in eleventh grade mathematics and science courses. Less than half of eleventh graders enrolled in mathematics and science courses compared to middle class and affluent schools who had 60 to 100 percent of students enroll in those classes (Spade et al., 1997). Additionally, school personnel were less active in working class communities in the course selection process and more weight was given to students' and parents' choice. In middle- and upper-class schools, objective measures, such as test scores, prior performance, and teachers' recommendations were used more in the process to enroll students in certain courses (Spade et al., 1997).

Riehl, Pallas, and Nariello (1999) found in their study that the processes to determine which students could access honors or AP courses was a socially constructed process within the schools. The researchers studied course scheduling in five urban high schools. In each school, minority students' enrollment was over 90 percent with many students eligible for free and reduced lunch. Staff members (n=78) were interviewed in each school who were knowledgeable about the school scheduling process (Riehl et al.,

1999). The data showed that each high school did have some courses just for high achieving students and designating students for challenging classes was different depending on the school. In one school, Advanced Placement English classes were created by teacher recommendation. In another school, staff first advertised AP classes among teachers and students. They then worked from a list of interested students, gave a reading and writing assignment to the students, asked for two teacher recommendations and interviewed them. A committee then voted on who would be admitted into the classes. A staff member explained that some students were afraid of doing all that work to enter a class. Another relayed that honors classes had become a way to sort and separate students. Many of the schools also relied on standardized test performance to separate students into classes. This system persisted even though some test scores were out of date. For example, in one school the incoming ninth grade students were assigned classes based on their seventh-grade test scores. In all schools, the researchers ultimately found that the scheduling process was really a "socially constructed ritual" (Riehl et al., 1999, p. 145). While this problem might not always be negative, it can bring inherent bias because of institutionalized beliefs and understandings. As institutional norms change, the routines for curricular choices might change as well. One way to revise a socially constructed or normative base for course scheduling might be through detracking initiatives (Riehl et al., 1999).

Kelly (2007) in North Carolina found that course placement criteria in North Carolina was largely based on tracking with little mobility between tracks. The data used in the study was from curriculum guides from North Carolina high schools in the 1997-98 school year. Data from 92 (n=92) of the 351 public high schools were used. The results

showed that all public high schools in North Carolina almost all schools had multiple English and math courses for students of different ability levels. In some schools, social studies and science were just as selective as math and English and in others, they were the heterogeneously grouped classes. To determine who could enroll in certain classes, schools often referred to their track assignment process as a student/parent choice system, however this was not the reality. Students were allowed to choose which classes they could enroll in if they were eligible. Eligibility would be based on whatever the requirements each school chose. For example, one high school claimed to have an open enrollment policy but they had a number of prerequisite requirements such as meeting a certain grade in previous courses, scoring above a certain cutoff point on standardized tests, prior course taking, and teacher recommendations (Kelly, 2007). Additionally, mobility between tracks was often difficult, especially if moving up to a higher track. Students, in thirteen of the schools, were not allowed to switch into the honors track if in a lower track. However, students could move between the lower two tracks. There were also subjective placement requirements for students. For example, 43 percent of schools required teacher recommendations in English class placements and 50 percent did with science placements. Some schools required recommendations for only the students switching tracks. In one of the schools, students chose course sequences based on students' occupational goals in 8th grade. At this time, many students did not realize the importance of course taking on their future opportunities. This polarization of students into different levels based on subjective course placement strategies reduced their opportunities in their education while detracking efforts opens these opportunities to all (Kelly, 2007).

Another study (Mickelson & Everett, 2008) focused on the occupational tracking system instituted in the North Carolina schools specifically. The results showed that the social stratification in society was reproduced in the schools particularly in predominantly white schools because of the tracks students chose to enroll in. Mickelson and Everett (2008) explained that African Americans, in North Carolina, who attended diverse schools were less likely to enroll in AP classes than African Americans who attended primarily African American schools. In response to the state's new standards at this time, schools implemented the High School Courses of Study Framework (COS) and with that came a form of neotracking. The COS required 8th graders to choose one of the three courses before entering high school: Career Preparation, College/Tech Preparation, or College/University Preparation. Exceptional or gifted children took Occupations, a fourth COS. Students were allowed to switch Courses of Study during high school. Students were also allowed to take whatever AP or IB level course they wished to take as long as they completed the prerequisites. The findings showed that the variations in enrollments reflected the race, ethnic, and social class of the students in North Carolina (Mickelson & Everett, 2008). Wealthier students from more affluent schools, were significantly more likely to choose the top COS than those who came from poorer schools. Additionally, COS enrollments were related to the race of the students. For example, 85.6% of Asians compared to 56.9% of American Indian students enrolled in the College/University Prep COS. Black, Latinx, and American Indian students were more likely to choose Career Prep than White or Asian students; Asians were the least likely to choose Career Prep. Black students were the most likely to choose Career Prep out of any ethnic group, followed by Latinx, and then Whites. Although the majority of

students in every racial group selected the College/University Prep COS, there were still statistically significant differences in choice based on race. Furthermore, Mickelson and Everett (2008) found that the more racially isolated their minority high school was, the more likely they were to be enrolled in an academic COS. For example, 73.7% of Blacks in racially isolated White high schools, 81.6% in racially balanced schools, and 87.9% in racially imbalanced Black schools enrolled in the College/University Prep COS. Also, an average of 12.4% of Blacks in racially isolated White high schools, 5.7% in racially balanced schools, and 4.2% in racially imbalanced Black schools enrolled in Career Prep COS. In one affluent area of North Carolina, Black students were also less likely of all ethnic groups to enroll in the College/University Prep COS (36.5%). Only .3% of Whites, 3.1% of Asians, and no Hispanics were enrolled in the same COS at that particular school. Thus students, and the curriculum they chose to study, were separated by race especially in diverse schools because of tracking (Mickelson & Everett, 2008).

Lewis and Cheng (2006) also found that socioeconomically disadvantaged students were more likely to be in non-college tracks, that SES predicted tracks in schools, and that SES predicted principal's expectations of students' futures. Principals responded by survey that asked how many tracks their schools had. The survey also showed that students' requests (M= 4.06) on a scale of 1 to 5 were the highest rated criteria to enroll in classes. Other criteria that rated higher than 3.0 were special learning needs, parental requests, and guidance counselor assessments. To examine the determinants of track structure, a multinomial logistic regression analysis showed that when the college category was compared to the vocational category, free and reduced lunch percentage made a statistically significant difference. So, whether the vocational

track or the college track was predominant in a school was determined by the number of students receiving free and reduced lunch (Lewis & Cheng, 2006). Principals (*n*=455) also had to indicate percentages of students in their school that they expected to follow different paths post high school. The predominant expectation of principals was four-year college at 60.2 percent followed by community college at 19.3 percent. Then, a multinomial logistic regression was used to determine what independent variables predicted the dominant expectation. The independent variables were two SES measures, and the race/ethnicity measures. When the four-year college category was compared with the work category, income and free and reduced-price lunch was the significant difference between them. When comparing the community college category to work category, free and reduced-price lunch was again the source of significant difference. Therefore, free, and reduced-price lunch status was a strong predictor of principals predicting what students would do after high school and SES also predicted the dominant track in schools separating students based on SES (Lewis & Cheng, 2006).

Another qualitative study focused on the experiences and low enrollment of African American students in challenging courses (Lofton, 2019). This study included twenty-six African American parents (*n*=26), many of whom attended the same school district in a racially diverse school, came from a segregated high-poverty African American community, and experienced lower track placement. In the middle school studied, Americans made up 14% and Whites about 65% of the schools' 900 students. African Americans scored pretty well on the school report card: 77% had 3s and 4s in English, 78% had 3s and 4s in Math, and 63% did the same in science on the standardized test in which 4 was the highest score. The middle school experienced a

battle when detracking students. While administrators supported detracking, many wealthy, White parents opposed it. The results of this fight led to a compromise in which the sixth grade was detracked, and seventh and eighth grade math and science classes were self-selected (Lofton, 2019). Despite the choice policy, test score results, observations, and interviews with school guidance counselors, data from the middle school office on schedules showed that African Americans were placed in special education, remedial, or standard classes in the "self-selected" courses. Most advanced math and science classes were composed of all White students (Lofton, 2019). African American parents and their children had experienced lower track coursework. It was found that tracking in this school was not based on ability but rather on the desire to segregate African American students (Lofton, 2019). Moreover, African American students had to face the challenges of systemic racism, such as housing struggles, that other students in Advanced Placement class did not encounter which impacted their grades and therefore their placements (Lofton, 2019). The researcher explained that the choice policy caused a "duplicity of equality" for African Americans because the school seemed to offer policies that equalized students but actually still just sorted African Americans into lower-level classes (Lofton, 2019, p. 25). Therefore, social stratification we merely perpetuated over time unlike the fully detracked classes which broke down this separation (Lofton, 2019).

Moreover, another study by Yonezawa, Wells, and Serna (2002) found that course self- selection resulted in classes being segregated by race and class with low income and minority students in the lowest level classes. Tracks had created high and low track cultures and the students' identities were often reflected in their choice for a certain

academic track. Leaving a certain track was not easy for students because they had always identified with that track and thought of themselves in terms of "low track" or "slow" or "high track" and "gifted" (Yonezawa, et al., 2002, p. 52-53). The educational system had created these identities because of the labeling of students at a very young age. Students labeled gifted in elementary school felt entitled and saw high track classes as their correct placement. This labeling was constantly reinforced and merely followed the students throughout their schooling because they saw it as their educational progression (Yonezawa, et al., 2002).

Because of tracking policies and gatekeeping tactics both between and within schools, equity gaps still exist. One study by Kettler and Hurst (2007) examined the equity gaps in a sample of suburban public high school and how these gaps changed over time. It also researched the relationship between school factors and the magnitude of the ethnicity gaps in suburban schools in 2011. The study focused on Black, Hispanic, and White student populations and records of AP and/or IB enrollment from 117 suburban high schools (n = 117) in Texas. The researchers used a multiple regression analysis to determine the ethnicity gap in AP/IB participation. The results of the study showed that the mean participation rates in AP/IB testing at the two-time intervals, 2001 and 2011 was statistically significant (t_{232}=5.66, p<.001, Cohen's d =0.74 (Kettler & Hurst, 2017). The Black student participation rate was also statistically significant (t_{182}= 6.13, p<.001, Cohen's d =0.91) as well as the Hispanic participation rate (t_{230}= 5.81, p<.001, Cohen's d =0.76) and White participation (t_{221}= 6.11, p<.001, Cohen's d =0.82). All three groups significantly increased enrollment over time in the schools. In 2001, the analysis showed significant participation gaps between Hispanic and White students and Black and White

students who participated in AP/IB testing. For both Black and Hispanic students, these ethnicity gaps did not significantly increase or decrease over the eleven-year period (Kettler & Hurst, 2017). Regression analysis further predicted what factors may have been associated with those ethnicity gaps. The school achievement index was the strongest predictor of the Black/White participation gap however the proportion of the school faculty that was minority was a strong predictor for the ethnicity gap for Hispanic students. As the proportion of minority faculty increased for the school, the AP/IB participation gap for Hispanic students decreased. Ethnicity gaps in these school districts remained the same even though more students participated in the programs (Kettler & Hurst, 2017). Schools must evaluate and implement detracking policies focused on subgroups to close the equity gaps existing.

Students' Perspectives. In fact, detracking greatly helps all students socially, emotionally, and academically according to one study by Yonezawa and Jones (2006). The study examined students' perceptions of tracking and detracking through a qualitative based study. Most students felt that the tracking system was inequitable and made them uncomfortable. Very few knew how tracking placements occurred. Many were frustrated that schools used test scores to define them and their eligibility into classes. They also felt that weaker students received less caring and less knowledgeable teachers. One student explained that his/her history teacher focused more on her AP classes than her lower-level classes and sometimes would show up to school just for her AP classes making the lower track students feel inadequate and uncared for. At one third of the schools, only a few backed tracking and the others promoted detracking. Importantly, the students also argued that detracking success would take more than policy

change. They claimed that it would take teachers to believe in them to really make a positive difference. Many students were concerned that teachers did not have high expectations for all students. A student, in one group, stated that his calculus teacher degraded the remedial algebra students by making fun of them (Yonezawa & Jones, 2006). Students wanted more challenging courses regardless of their track level. Students from lower tracked classes said they often sat through boring, repetitive material and rote tasks. They felt they were not challenged by the curriculum or their teachers. Therefore, detracking policies but also a shift in mindset would lead to student success (Yonezawa & Jones, 2006).

Similarly, in another study, Boaler, Wiliam, and Brown (2000) argued that students in lower tracks received less curriculum and were taught in less imaginative ways. There were also lower teacher expectations in lower tracks and less time spent on higher level, critical thinking skills. The researchers studied students' attitudes and achievement in mathematics and the influence of ability grouping on these experiences in six schools in the United Kingdom. The students (n=943) moved from mixed ability grouping in grade eight to tracked classes by ability in grade nine in four of the six schools. What resulted was "curriculum polarization" (Boaler, Wiliam, & Brown, 2000, p. 631) because the top students were expected to move at a pace that many students could not keep up with, but it was even worse for the lower students who experienced less opportunities to learn. For the lower students, student interviews described high teacher turnover and change. Students also said that their work was too easy and low level. Even worse, teachers ignored students asking their teachers for more difficult work. Some students were just given assignments to purely copy answers off the board and

others finished their work within the first five minutes of class with nothing else to do. Students from the lower groups explained that there was little movement of students to higher tracks. Students felt trapped in lower tracks because they were not taught the skills and content to pass the test for the higher track classes. This problem of low-level work seemed to stem from teachers' perceptions that the work matched students' abilities and also that students in lower groups have the same capabilities and learning styles as each other (Boaler et al., 2000).

Impact on Curriculum and Instruction. Later, Boaler (2011) completed two longitudinal research studies to examine the highly successful teacher practices when working with heterogeneous groups and another third study in which the researcher and graduate students taught four mixed ability classes for sixth and seventh graders in order to put the successful teaching practices into place. From all three studies, according to Boaler (2011), there were a few similar pedagogical practices that led to successful detracked classrooms fixing many of the issues that students in tracked environments faced in the previous study (Boaler et al., 2000). The teachers who were successful in detracked environments did the following: students worked on mathematics tasks that were appropriate for different achievement levels, students were taught to respect each other and work together, and teachers gave messages that all students could learn because of a growth mindset (Boaler, 2011, p. 11-12). Teachers achieved the students working at all different levels in English by giving "open, exploratory tasks" that students could answer in all different ways (Boaler, 2011, p. 11). Teachers always gave messages that all were capable and also taught students how to listen to each other, how to be respectful, and how to work together. They described learning as a process and that high

achievement was due to effort, not ability. Teachers also made sure to praise low achievement students publicly about their contributions to the class (Boaler, 2011). These teaching practices led to successful and effective detracked classrooms in terms of social and emotional skills and academics (Boaler, 2011).

Impact on Academic Outcomes. In terms of achievement, research has revealed that when advanced level coursework was encouraged and extended to minority students and students with lower socioeconomic status and when teachers receive the training and resources needed, those students succeeded just as other students. In fact, students in more challenging courses were more engaged with school which led to higher graduation rates, fewer cases of suspension, and fewer absences. Therefore, the reasons these opportunity and achievement gaps exist is because of policy decisions of adults and not the academic abilities of the students (Patrick, Socol, & Morgan, 2020). The elimination of tracking, ability grouping, leveling or any form of educational stratification has been found both effective and beneficial to all students (Burris & Garrity, 2008). Both the College Board and the International Baccalaureate Organization have promoted open enrollment or detracking in their programs to increase equity but importantly this didn't impact high achievers and improved the scores of lower achieving students (Kettler & Hurst, 2017).

In one study, a diverse, suburban high school implemented a policy of detracking or unrestricted access to the IB program and the results showed achievement successes for students (Atteberry, Lacour, Burris, Welner, & Murphy, 2019). The reform movement detracked all courses in Grades 6 through 10 and then also opened IB enrollment in Grades 11 and 12. The data analyzed were from an 18-year period in which the IB

program expanded from serving a few, top students to serving a majority of the school's students (Atteberry et al., 2019). By the year 2011, students experienced all subjects completely detracked from grades 6 through 12. In order to measure achievement prior to IB enrollment, the researchers used students' tenth grade scores on the verbal and mathematics PSAT test and categorized them into low, middle, and high performing as they entered IB grades. The analysis then looked at achievement on the IB assessments. By using an interrupted time series model, the researchers could estimate the impact of detracking on IB scores. The analyses were run for each detracking phase as well as for low, middle, and high achieving students (Atteberry et al., 2019). The results of the study were favorable toward detracking. The overall participation in both ELA and Math increased from about 20-30% of students having an IB score to 70-80% (Atteberry et al., 2019). Moreover, more students with lower PSAT scores began taking IB classes. Those same students would have been denied entry to these classes before the detracking efforts. In terms of achievement, for IB math tests, student performance increased during the beginning phases of detracking but evened out at the end. However, the average IB score did not significantly change over the years. There were no statistically significant differences in trends over time. So, detracking was not associated with lower mean math IB performance in any group even though there was increased participation (Atteberry et al., 2019). For English, in Phase I and II of detracking, trends in IB scores were more varied across prior achievement groups however, the trends for all groups were mostly flat during the final phase of detracking. So, the students in Phase III scored similarly to the students in Phase I even though there was an increase in students taking the IB courses over time. There was no statistically significant difference in mean IB scores or

trends in IB scores across the phases of detracking students (Atteberry et al., 2019). Importantly, many of these results were not statistically significant, showing that even when students with low PSAT scores were added to the enrollment in IB classes, the achievement on IB assessments did not change (Atteberry et al., 2019).

Another successful example of detracking occurred at the Preuss School on the University of California San Diego campus. The University of California San Diego established a charter middle/high school to prepare students from low-income families to prepare them for college. This study by Alvarez and Mehan (2006) argued that detracking students and presenting students with a rigorous curriculum along with academic support can push students from low-income backgrounds to college enrollment. The curriculum was purely focused on college prep with multiple supports to help. The structural and cultural change brought great success to this detracked school. Eighty percent of the graduating class (n=55) attended a four-year college and twenty percent attended community colleges with a transfer to UC colleges in two years. In terms of achievement, the Preuss School students scored above the 50th percentile in reading on the CAT/6 reading test in 2002-2003. Ninety percent of the grading class of 2004 passed both portions of the California High School Exit Exam by March 2003 and 92% of the 2005 class passed both portions by Match 2003. Of the students in grades 8 to 11, 327 took AP exams and 37% scored a 3 or higher. Lastly, the Preuss School had the highest API Scores in San Diego County for schools with greater than 80% of students eligible for free and reduced lunch (Alvarez & Mehan, 2006). The Preuss School was able to improve academic achievement for students of low socioeconomic backgrounds. Through

detracking and pushing rigorous classes for all, expectations of students were raised and therefore, many students aimed and achieved those expectations.

Impact on Students' Future Decisions. The changing educational requirements for post-secondary careers emphasize the importance of a challenging high school curriculum for students. Between 1992 and 2016, the number of people who had earned a bachelor's degree increased from 18 to 25 percent and those with advanced degrees (master's, professional, and doctoral) had increased from 9 to 14 percent. In 2016, two thirds of the labor force had some college experience. Moreover, the number of occupations requiring college degrees has increased. Of the 30 fastest growing occupations, 18 of those require more than a high school diploma according to the Bureau of Labor Statistics (Khine, 2019). One approach to strengthen high school students' college readiness is to increase their participation in AP courses; selective colleges also take students' enrollment in AP courses into account for admission into college (Dougherty, Mellor & Jian, 2006). Therefore, the trajectory of students' programs in their secondary schooling experiences can impact their future success.

Dougherty, Mellor, and Jian (2006) completed a study that examined if students in AP courses graduate from college at higher rates than non-AP students, if the same is true when controlling for students' observed characteristics and the characteristics of their schools, and if high schools with a higher percentage of students in AP courses have higher graduation rates of their students who attend college controlling for measured characteristics of the school. To answer these research questions, the researchers followed a statewide cohort of 54,566 (n= 54,566) 8th graders in Texas who graduated from high school in 1998 and enrolled in a Texas public college or university within

twelve months of graduation. Students were divided into four groups: passed AP exam, took but did not pass AP exam, took AP course but not AP exam, and took no AP course or exam (Dougherty, et al., 2006). The results indicated that the willingness of students to enroll in AP courses and take an AP exam were more likely to graduate from college. For example, low-income students who took and passed at least one AP exam had a 39-percentage point higher college graduation rate (46% vs. 7%) than low-income students who did not take any AP course or exam. African American students who took and passed at least one AP exam also had a higher college graduation rate of 53% compared to those who took no AP course or exam whose graduation rate was 10%. Hispanic students saw the same results with an even larger gap; those who passed an AP exam had a graduation rate of 54% and those who did not had an 8% graduation rate. The study also found that there was a statistically significant relationship between AP exam passing and college graduation for all groups except African Americans. Only 61 African Americans in the cohort passed at least one AP exam. This data showed that the percentage of students who took and passed AP exams was an indicator of whether the school was preparing students to graduate from college. It also showed that although the number of low income and minority students taking AP courses seemed to have increased, the percentage of students who passed those AP exams was still low. There was still a major gap in preparing low-income students for the exams and college. Consequently, schools need to add to the number of AP courses offered but also prepare students of all backgrounds and levels to take the courses. This preparation and programming should start in early grades so that they can do the college level work while in high school (Dougherty et al.,2006). Due to the increased need to attend college and

the relationship between AP coursework and college success, schools must look at detracking initiatives to increase both the enrollment and achievement of students in AP classes.

Relationship Between Prior Research and Present Study

According to the literature, some proponents of tracking argue that it leads to increased achievement scores for all groups of students, particularly the top students (Duflo, Dupas, & Kramer, 2011; Epstein & MacIver, 1992, Figlio & Page, 2002; Loveless, 2016). However, many different variables can confound these results such as race, ethnicity, and SES. Moreover, there is no causal link that tracking helped teachers in their pedagogies tailor their instruction to homogenous groups leading to increased student achievement (Duflo et al., 2011). According to the literature, there are also mixed results concerning students' emotional impact from tracking (Kulik & Kulik, 1992; Wiatrowski, 1982; Yonezawa & Jones, 2006; Yonezawa, Wells, & Serna, 2002) as well as tracking's effects on students' future academic success (Clark, 2010; Dougherty, Mellor, & Jian, 2006). Through a utility lens, tracking might be beneficial but in regard to equity, it is unjust. If the end result is to focus on equity and to help the underrepresented populations in Advanced Placement classes, detracking policies in secondary, public schools provide all students access and opportunities (Burris & Garrity, 2008). It is not acceptable to raise the overall outcome of test scores at the expense of some students. Detracking studies improve curriculum and instruction for students, especially those who were in the lower tracks, (Yonezawa & Jones, 2006; Boaler, 2000; Boaler, 2011) the social and emotional well-being of students (Yonezawa, Wells, & Serna, 2002), and also leads to improved academic achievement due to higher expectations for all (Atteberry et

al., 2019; Alvarez & Mehan, 2006). The present study looked at trends before a detracking policy, during the policy change, and after to determine the impact on achievement results on the whole student population taking AP classes and on certain demographic groups. This study determined if the increased access and opportunity had a positive impact on enrollment and achievement of Advanced Placement students and if a policy such as Challenge by Choice can effectively close the opportunity and achievement gaps pervasive in schools.

CHAPTER 3

The purpose of this study was to determine how a Challenge by Choice detracking policy impacted both students' enrollment and achievement in AP courses. Prior to this policy change in 2011, the school district enrolled students into AP courses only if they had a certain grade point average and a teacher referral. This study examined both how this change in policy in 2011 impacted all student groups and different demographic groups by comparing trends before the policy, during the policy change, and post policy years.

Research Questions and Hypotheses

This study was guided by two primary research questions that explored how a Challenge by Choice policy affected students' enrollment in AP courses and their subsequent AP scores.

Question 1. How did the introduction of AP Challenge by Choice change student's AP course-taking behavior?

A. How did the introduction of the AP Challenge by Choice policy change the percent of students taking at least one AP course during high school, overall or by student subgroup?

H_0: The Challenge by Choice policy did not change the percent of students taking at least one AP course during high school, overall or by student subgroup.

H_1: The Challenge by Choice policy did change the percent of students taking at least one AP course during high school, overall or by student subgroup.

B. How did the introduction of the AP Challenge by Choice policy change the average number of APs students elected to take during high school, overall or by student subgroup (White, minority, female, male, ELL, and students with IEPs)?

H_0: The Challenge by Choice policy did not change the average number of APs students elected to take during high school overall or by subgroup.

H_1: The Challenge by Choice policy did change the average number of APs students elected to take during high school overall or by subgroup.

Question 2. How did the introduction of Challenge by Choice change the trend in AP enrollment and test scores, overall and by student subgroup?

H_0: The enrollment and AP scores of students do not differ prior to and after Challenge by Choice was implemented overall or by subgroup.

H_1: The enrollment and AP scores of students do differ prior to and after Challenge by Choice was implemented overall or by subgroup.

Research Design and Data Analysis

A quantitative study of secondary data was used to determine the impact of the Challenge by Choice policy. First, I used a one-way ANOVA to compare the means between the groups that had no exposure, some exposure, and full exposure to the policy. The ANOVA test determined if the percentage of students taking at least one AP course and the average number of AP courses students elected to enroll in varied among these groups (Creswell & Guetterman, 2019). Second, I used an interrupted time series analysis

to measure the dependent variable (student enrollment in Aps, number of AP courses students elected to enroll in, and students' AP Scores) before and after the Challenge by Choice policy went into effect (Chiang, Jhangiani, & Price, 2015). It was appropriate in this instance because the researcher obtained multiple data for a pretest measured for a period of time, data from the time of the intervention, and then data from posttests for a period of time to determine the statistical significance of the intervention (Creswell & Guetterman, 2019).

Data Collection

After receiving IRB approval and permission from the district superintendent to collect data, the researcher asked for students who took AP exams for a span of fifteen years including time before the Challenge by Choice Policy went into effect (2005-2010), during its implementation (2011) and after the Challenge by Choice Policy was implemented (2012-2019). Data on the students who did not take the exam during these time periods was also requested. The researcher accessed this information from the Director of Guidance and Data Management. After receiving consent, the researcher received all information from the Director of Guidance and Data Management. Participants in the data collection remained anonymous and confidential.

After the Director of Guidance and Data Management accessed the information from Naviance and PowerSchool, it was sent to the researcher. The information included exam grades (1-5) on the AP Exams and student demographic information. No personal data, such as name, address, or phone number, was collected to preserve anonymity. The information was sent to the researcher as an export from the Naviance and PowerSchool systems. This was then exported to a Microsoft Excel Spreadsheet. Values in both

spreadsheets were checked, so there was no error in the transfer of the information. Afterwards, the Excel document was transferred to SPSS, and once again data was checked for accuracy.

Sample and Population

The population for this study was comprised of students from the years 2005-2019. The data was collected from a small, suburban school district in Suffolk County, NY. There are approximately 1,000 students at the high school level. Demographic information for the student population in each 9th Grade cohort used in the study are found in Table 1.

Table 1
Demographic Information of Student Population

Year	Total	Minority	White	ELL	Female	Male	IEPs
2005	277	34	243	4	141	136	27
2006	273	38	235	5	135	138	33
2007	280	48	232	3	134	146	28
2008	266	41	225	3	119	147	5
2009	289	53	236	11	152	137	26
2010	277	48	229	5	123	154	27
2011	288	57	231	7	135	153	20
2012	291	74	217	6	121	170	32
2013	282	74	208	6	140	142	28
2014	254	60	194	11	110	144	26
2015	293	80	213	15	137	156	31
2016	272	79	193	15	128	144	29

Note. Year indicates Spring of school year, e.g., 2005 indicates the 2004-2005 school year.

Instruments

As this is an *ex post facto* study, the researcher did not use an instrument to collect any student information; however, all test scores were taken from the Advanced Placement examinations in a variety of subjects. The Advanced Placement program

allows high school students to receive college credit for courses taken in high school. The program has expanded extensively through the years. The courses end with nationally standardized examinations and are administered in May of every year. The examinations are developed by committees of secondary teachers and college professors in each subject. Tests are about three hours in length and usually consist of multiple-choice questions and essays or problem-solving questions with constructed responses. For each of the multiple component AP examinations, students receive a weighted sum of the component scores depending on the importance of each part. The AP grades are then transferred to a five-point scale by dividing the composite scores into five intervals based on cut scores for each program. Those cut points are created by the Chief Reader who uses equated multiple-choice scores as their reference (Bridgeman, Morgan, & Wang, 1996).

Reliability. Bridgeman, Morgan, and Wang (1996) found that the reliability of the multiple-choice sections and the composite scores for all examinations was high (.82) at the time of their study. For the interrater reliability, Bridgeman et al. (1996) found that it could not be routinely predicted because the responses are read by a single reader however there have been special reader reliability studies that compared the correlations of the first and second readings to determine reader reliability. There was much variation in the reliabilities of free response sections. For example, the reliability for Computer Science A was .88 but for European History, it was .49. Free response scores in foreign languages, science, and mathematics were more reliable than essay scores in the humanities. Interrater reliabilities also varied ranging from .98 in Computer Science A to .67 in U.S. History. Rater inconsistency seemed to be the issue in these cases rather than

raters scoring too high or low on sections. Rater training and adding topics for free responses improved the reliability for both (Bridgeman et al., 1996).

Lastly, Bridgeman et al. (1996) determined the reliability of the AP grades themselves. The grades are given on a 1-5 scale. The algorithm was originally devised by Livingston and Lewis (1991) and then revised by RELCLAS-COMP Version 4 by 1996 to test the reliability of AP Composite scores into the grading scale and to determine the accuracy of that grade. At the time of this study, RECLAS also estimated the proportion of students who were above or below each of the four cut scores (Bridgeman et al., 1996). Although the study found much variation in the students receiving AP Grades on a 1-5 scale, the high percent of students (87% to 93%) who correctly received a passing grade (3 or above) or a failing grade (2 or below) made the exams reliable (87% to 93%) (Bridegman et al., 1996).

Validity. The validity of the AP test has become increasingly pertinent as some colleges give AP course credit to students while others refuse to recognize AP classes as equivalents to college courses ("More Colleges Stop Giving," 2019). One study by Hargrove, Godin, and Dodd (2008) did a statewide study on students' performance on several college outcomes such as first and fourth year GPAs, first- and fourth-year credit hours earned, and four-year graduation status differentiated by the AP experiences students had in high school. The students who took AP courses and exams outperformed the AP course only and other course group students. This occurred for both first- and fourth-year GPAs for the 1999-2002 and 1999-2001 cohorts of students in AP English Language and Composition, AP Literature and Composition, AP Calculus AB (except for fourth year GPA in 1999), AP Biology (except in 2000), AP Chemistry (in 2001 and

2002 only), and AP U.S. History (also in 1998 and 1999) analyses. The AP course and exam group also earned higher mean GPAs than the other course only students in AP English Language and Composition (except in 1998 first year GPA), AP English Literature and Composition, AP Spanish Language (except in 2000). Furthermore, AP course and exam students graduated at significantly higher rates within four years than students in the AP course only or other course groups in almost all the subjects for the 1998-2001 cohorts (Hargrove, Godin, & Dodd, 2008).

Another study by Murphy and Dodd (2009) found that AP students as a group consistently outperformed non-AP groups of students. Across the ten subjects, students who had taken an AP exam took more credit hours in their first year, in the subject area, and in college overall than non-AP students. They also had higher GPAs in the subject area related to the AP exam taken. For five of the ten AP exam subject areas, the AP students had significantly higher first year college GPAs. The study also found that students who had AP credits significantly outperformed the students that did not have AP credit. Effect sizes for this were moderate across multiple subjects for cumulative college credit hours taken and first year GPAs (Murphy & Dodd, 2009).

Mattern, Shaw, and Xiong (2009) also found that higher AP performance on the English Language, Biology, Calculus AB, and U.S. History Exams corresponded to higher first year GPAs, higher second year retention rates, and attendance at selective colleges/universities. Even when prior academic performance was controlled for (SAT scores and HS GPAs), significant group differences existed for the AP performance group across four of the most commonly taken AP exams. Students who receive a 3, 4, or 5 on one of the AP exams outperformed students who received lower AP scores and

students who did not take an AP exam in regard to first year GPA, retention, and institution selectivity. Similarly, students who took an AP exam but scored a 1 or a 2 and attended more selective institutions were more likely to return their second year than the non-AP group (Mattern, Shaw, & Xiong, 2009).

Mattern, Marini and Shaw (2013) in their study found positive support for the AP program and its validity in measuring college success even when controlling for previous academic performance and demographic variables. The researchers found that participation in the AP program and the completion of one or more AP exams was related to an increased likelihood that the student would graduate in four years or fewer. Graduation rates were 58% and 55% for AP exam takers compared to 38% and 34% for non-AP students in the two samples of students. The results also showed that students who earned higher scores were more likely to graduate in four years or fewer compared to those who earned lower scores. The four-year graduation rates were 37% and 34% for students with an average AP score of 1 and 71% and 73% for students with an average score of 5. The first sample of students was based on four years of data collected on a sample of first time, first year students who entered college in the fall of 2007. The sample contained 112,108 students (n=112, 108) from 77 colleges and universities. Students had valid PSAT scores and demographic information. The second sample consisted of 1,297,950 students (n=1,297,950) who enrolled in a four-year institution after high school. This sample of students also all had a valid PSAT score and demographic information (Mattern, Marini, & Shaw, 2013).

Still other studies did not find that taking more AP classes led to increased college readiness. Warne (2017) contends that previous studies on this subject are extremely

biased if covariates are not controlled since most AP students are more likely to be White, from suburban areas, and have better academic preparation than non-AP students. Warne cautioned against implementing more AP courses as an answer to increased college readiness and oversimplifying the impact of the program. Instead, specific students or subpopulations and courses must be studied (Warne, 2017).

Summary

Overall, the reliability of AP exam composite scores was overall very high (.82). Although the free response sections and essay tasks caused some variation in reliability, studies consistently strive to improve interrater and topic reliability. Moreover, the percent of students who correctly received a score of 3 or higher or a 2 or below was very high (87% to 93%) making the tests reliable (Bridgeman et al., 1996). In terms of validity, taking AP courses and exams were positively associated with success in college and graduation rates, in many AP course subjects (Hargrove, Godin, & Dodd, 2008; Hargrove, Godin, & Dodd, 2008). Even when prior academic performance was controlled, AP students had higher college success and retention rates (Mattern, Shaw, & Xiong, 2009; Shaw, Marini, & Mattern, 2013).

Data Analysis

Research Question 1A

Descriptive statistics were used to calculate the percentage of students taking at least one AP course by 9th grade entry cohort. Cohorts were divided into three categories: those unaffected by the policy (having graduated prior to policy implementation), those who experienced the policy change during HS (those who were in 10th, 11th or 12th) grade at the time of policy implementation, and those who always were exposed to the policy

(entering 9th grade during or after policy start). A one-way between-subjects ANOVA was used to compare the percent of students taking at least one AP by group to determine whether there were significant gains in enrollment in APs as a result of the policy.

Descriptive statistics were also used to summarize any changes in subject-area AP courses taken by students by cohort, identifying which AP course enrollments increased the most post-policy.

Research Question 1B

Descriptive statistics were used to calculate the average number of AP courses students elected to take by 9th grade entry cohort. Cohorts were again divided into three categories: those unaffected by the policy (having graduated prior to policy implementation), those who experienced the policy change during HS (those who were in 10th, 11th or 12th) grade at the time of policy implementation, and those who always were exposed to the policy (entering 9th grade during or after policy start). A one-way between-subjects ANOVA was used to compare the average number of AP courses students elected to take to determine whether there were significant gains as a result of the policy.

Research Question 2

The following interrupted time series model was estimated to determine if the percent of students taking at least one AP course per year increased with the policy change.

$$AP_i = \beta_0 + \beta_1(t_i - t^*) + \beta_2(P_t) + \beta_3(P_t)\,(t_i - t^*) + e_t$$

In this model, AP_i is the percent of students taking at least one AP course per year; t is the year (e.g., 2010, 2011); t^* is the year in which the policy was adopted; and P_t is an

indicator that the observation is from the post-period (after policy implementation). Therefore, β_1 is the trend in percent of students taking at least one AP per year prior to the implementation of the policy; β_2 is the "jump" in the percent of students taking at least one AP course per year that occurred at the time of implementation of the policy change; and, β_3 is the "shift" in the trend in the percent of students taking at least one AP course after implementation. Coefficients were considered significant at the $p < 0.05$ level. To understand whether the impact differed by subgroup, the model was estimated separately for each subgroup and the coefficients from the model were compared.

The following interrupted time series model was estimated to determine if the average number of APs a student takes per year increased with the policy change.

$$NumAP_i = \beta_0 + \beta_1(t_i - t^*) + \beta_2(P_t) + \beta_3(P_t)\,(t_i - t^*) + e_t$$

In this model, $NumAP_i$ is the number of APs each student took during each year; t is the year (e.g., 2010, 2011); t^* is the year in which the policy was adopted; and P_t is an indicator that the observation is from the post-period (after policy implementation). Therefore, β_1 is the trend in number of APs taken per year prior to the implementation of the policy; β_2 is the "jump" in number of APs taken per year that occurred at the time of implementation of the policy change; and, β_3 is the "shift" in the trend in number of APs taken after implementation. Coefficients were considered significant at the $p < 0.05$ level. To understand whether the impact differed by subgroup, the model was estimated separately for each subgroup and the coefficients from the model were compared.

By exam subject-area, the following interrupted time series model was estimated to determine if the AP scores of students on the exam differ pre- and post- policy change.

$$APScore_i = \beta_0 + \beta_1\,(t - t^*) + \beta_2\,(P_t) + \beta_3\,[(t - t^*)P_t] + e_t$$

In this model, $APScore_i$ is the student's AP score (on a scale from 1 to 5); t is the year (e.g., 2010, 2011); t^* is the year in which the policy was adopted; and P_t is an indicator that the observation is from the post-period (after policy implementation). Therefore, β_1 is the trend in AP scores prior to the implementation of the policy; β_2 is the "jump" in AP scores that occurred at the time of implementation of the policy change; and, β_3 is the "shift" in the AP scores after implementation. Coefficients were considered significant at the $p < 0.05$ level. To understand whether the impact differed by subgroup, the model was estimated separately for each subgroup and the coefficients from the model were compared.

CHAPTER 4

This chapter contains the analytic findings for the two research questions and subsets of research questions stated in the previous chapters. Prior to discussing the findings by question, I present a descriptive overview of AP course-taking in the school.

Descriptive Statistics

Table 2 shows the percentage of students taking at least one AP course by 9th grade entry cohort from 2005- 2016 broken down by subject. All subjects showed an overall increase between 2005 to 2016 in the percentage of students taking each AP course at the high school level. Table 3 shows the percentage of students taking at least one AP course by year and the average number of AP courses taken by students each year. In 2005, students took on average 1.65 AP classes compared to 2016 when students took on average 3.37 AP classes; the proportion of students taking APs also increased from .41 to .55.

Table 2

Percentage of Students Taking AP Courses by Subject

Cohort	ELC	CBC	*MAC*	*CAB*	*WHI*	*USH*	*BIO*
2005	32	5	24	13	23	19	9
2006	25	4	14	11	16	16	5
2007	25	4	16	13	18	14	6
2008	34	6	25	11	21	24	5
2009	29	8	27	17	22	27	6
2010	31	8	19	12	22	30	10
2011	32	8	32	11	25	26	8
2012	30	8	34	13	25	26	9
2013	39	13	41	12	34	34	10
2014	35	11	39	11	31	33	9
2015	29	6	29	15	28	33	11
2016	36	8	41	19	36	43	15

Table 3

Descriptives of AP Course Taking

Cohort	Percentage Taking At Least 1 AP	Average Number of APs
2005	41	1.65
2006	32	1.53
2007	30	1.63
2008	41	2.08
2009	40	2.33
2010	43	2.1
2011	46	2.24
2012	47	2.25
2013	51	2.86
2014	52	2.72
2015	47	2.46
2016	55	3.37

Research Question 1A

The first research question asked: How did the introduction of the AP Challenge by Choice policy change the percent of students taking at least one AP course during high school, overall or by student subgroup? To answer this research question, students were first divided into 3 groups and each student was assigned to a different policy group to determine if the introduction of the AP Challenge by Choice policy changed the percent

of students taking at least one AP class in high school overall or by student subgroup. Students in 9th grade cohort in the years 2005, 2006 and 2007 were never affected by the policy and thus were put into Policy Group 0. Students who belonged to a 9th grade cohort in 2008, 2009, and 2010 were impacted by the policy during at least one year while they were in high school and were put into Policy Group 1. Lastly, students who belonged to 9th grade cohorts in 2011, 2012, 2013, 2014, 2015, and 2016 were always exposed to the policy and put into Policy Group 2. Thus, three groups were created from the data: "no policy exposure," "some policy exposure," and "full policy exposure."

A one-way between-subjects ANOVA was used to determine whether there were any statistically significant differences among the means of the percent of students taking at least one AP course by policy subgroup. Before running statistical analysis, the data were screened and statistical assumptions were run. The Levene's Test of Variances was significant for the dependent variable "ap" which represented the percent of students taking at least one AP course during high school which indicated that there wasn't homogeneity of variances, $F(2,3,339) = 88.792$, $p < .001$.

Since the homogeneity of variances was significant, the researcher used the Welch's F-test to determine significance. The one-way ANOVA of the percent of students taking at least one AP course on the measure of cohort group revealed a statistically significant main effect, Welch's $F(2, 1792.272) = 28.310$, $p < .001$, indicating that not all cohort groups had the same mean percentage of students taking at least one AP course. Post hoc comparisons, using the Games-Howell post hoc procedure, were conducted to determine which cohort groups' means differed significantly. The Games-Howell post hoc test, as seen in Table 4, revealed that there was a statistically

significant mean difference in the percentage of AP course takers between the students who were not exposed to the policy at all and those who had some exposure to the policy (*MD* = -.068, *SE* = .024, *p* = .012), between those that had some exposure to the policy and full exposure to the policy (*MD* = -.084, *SE* = .021, *p* <.001), and between those who had no exposure to the policy and full exposure to the policy (*MD* =-.151, *SE* = .021, *p* <.001). Therefore, the percentage of students who took at least one AP in the group who had full policy exposure (Group 2) was fifteen percentage points more than that for students who had no policy exposure (Group 0). Due to the significant results, the null hypothesis was rejected; more students took APs as a result of the policy change.

To answer this question by student subgroup, the researcher replicated the analysis for the following subgroups: female students, male students, Asian students, Black students, Hispanic students, White students, non-Special Education students, Special Education students, non-ELL students, and ELL students.

Female students. The Levene's Test of Variances was again significant which indicated that there wasn't homogeneity of variances, $F(2,1572) = 5.042$, $p = .007$. The Welch's F test was used and was statistically significant, Welch's $F(2, 856.608) = 17.025$, $p < .001$. The Games-Howell post hoc test, as seen in Table 4, revealed that there was a statistically significant mean difference in percentage points between the females who had some exposure to the policy and full exposure to the policy (*MD* = -.103, *SE* = .031, *p* =.003), and between those who had no exposure to the policy and full exposure to the policy (*MD* =-.170, *SE* = .030, *p* <.001). Therefore, the percentage of females who took at least one AP in the group who had full policy exposure (Group 2) was seventeen percentage points more than the percentage of females who took at least one AP but had

no policy exposure (Group 0). Due to the significant results, the null hypothesis was again rejected.

Male students. The Levene's Test of Variances was significant which indicated that there wasn't homogeneity of variances, $F(2,1764) = 60.901$, $p < .001$. The Welch's F-test was used and revealed a statistically significant main effect, Welch's $F(2, 931.524) = 14.000$, $p < .001$. The Games-Howell post hoc test, as seen in Table 4, revealed that there was a marginally statistically significant mean difference in percentage points between the males who had some exposure to the policy and full exposure to the policy (MD = -.073, SE = .032, p =.056). However, there was a statistically significant different between those who had some exposure to the policy compared to those who had full exposure (MD = -.071, SE = .028, p =.032) and between those who had no exposure to the policy and full exposure to the policy (MD =-.144, SE = .027, p <.001). Therefore, the percentage of males who took at least one AP in the group who had full policy exposure (Group 2) was approximately 14 percentage points more than the percentage of males who took at least one AP but had no policy exposure (Group 0). Due to the significant results, the null hypothesis was rejected.

Asian students. The Levene's Test of Variances was not significant which indicated that there was homogeneity of variances, $F(2,75) = .200$, p =.819. However, for consistency with the other analyses, Welch's F is reported here and Games-Howell post-hoc tests were used. The one-way ANOVA of the percent of Asian students taking at least one AP course on the measure of cohort group revealed a statistically significant main effect, Welch's F (2, 46.719)= 5.791, p =.006. The Games Howell post hoc test, as seen in Table 4, revealed that there was a statistically significant mean difference in

percentage points between Asian students who were not exposed to the policy at all and those who had some exposure to the policy (*MD* = -.359, *SE* = .140, *p* =.037) and between those who had no exposure to the policy and full exposure to the policy (*MD* =-.402, *SE* = .124, *p* =.006). Therefore, the percentage of Asian students who took at least one AP in the group who had full policy exposure (Group 2) was approximately 40 percentage points more than the percentage of Asian students who took at least one AP but had no policy exposure (Group 0). Due to the significant results, the null hypothesis was rejected. The results were not significant between Asian students who had some exposure to the policy and those who had full exposure to the policy (*MD* = -.043, *SE* = .134, *p* =.945).

Black students. There were no significant mean differences for any groups of Black students during any point of the policy period. The Levene's Test of Variances was significant which indicated that there was not homogeneity of variances, $F(2,85) = 3.832$, $p = .026$. The one-way ANOVA of the percent of Black students taking at least one AP course on the measure of cohort group did not reveal a statistically significant main effect, Welch's F (2, 43.515)= .937, *p* =.399

Hispanic students. There were also no significant mean differences for any groups of Hispanic students during any point of the policy period. The Levene's Test of Variances was significant which indicated that there was not homogeneity of variances, $F(2,427) = 5.377$, $p = .005$. The one-way ANOVA of the percent of Hispanic students taking at least one AP course on the measure of cohort group did not reveal a statistically significant main effect, Welch's F (2, 135.543)= 1.273, *p* =.283.

White students. The Levene's Test of Variances was significant which indicated that there wasn't homogeneity of variances, $F(2, 2653) = 32.137$, $p<.001$. Welch's F-test was statistically significant, Welch's $F(2, 1511.929) = 33.215$, $p < .001$. The Games Howell post hoc test, as seen in Table 4, shows that there was a statistically significant mean difference in percentage points between White students who were not exposed to the policy at all and those who had full exposure to the policy (MD =-.180, SE = .023, p <.001) and also between the White students who had some exposure to the policy and those that had full exposure to the policy (MD =-.115, SE = .024, p <.001). There was also a significant difference between White students who had some exposure to the policy compared to those that had no exposure to the policy (MD =-.066, SE = .026, p =.032). Therefore, the percentage of White students who took at least one AP in the group who had some policy exposure (Group 1) was approximately 6 percentage points more than the percentage of White students who took at least one AP but had no policy exposure (Group 0). Due to the significant results, the null hypothesis was rejected.

Non-Special Education students. The Levene's Test of Variances was significant which indicated that there wasn't homogeneity of variances, $F(2,3027) = 26.315$, $p < .001$. The Welch's F-test was statistically significant, Welch's $F(2, 1622.740) = 28.971$, $p < .001$.The Games-Howell post hoc test, as seen in Table 4, revealed that there was a statistically significant mean difference between the non-special education students who had some exposure to the policy and full exposure to the policy (MD = -.104, SE = .022, p<.001), and between those who had no exposure to the policy and full exposure to the policy (MD =-.159, SE = .022, p <.001). Therefore, the percentage of non-special education students who took at least one AP in the group who had full policy

exposure (Group 2) was ten percentage points more than the percentage of non-special education students who took at least one AP but had no policy exposure (Group 0). Due to the significant results, the null hypothesis was rejected.

Special education students. The ANOVA could not be calculated for this subgroup due to small sample size.

Non-ELL students. The Levene's Test of Variances was significant which indicated that there was not homogeneity of variances, $F(2,3,248) = 80.456$, $p < .001$. The Welch's F-Test was statistically significant, Welch's $F(2, 1758.930) = 32.807$, $p < .001$.The Games Howell post hoc test revealed that there was a statistically significant mean difference between students who were not exposed to the policy at all and those who had full exposure to the policy ($MD = -.165$, $SE = .021$, $p<.001$) between students who were not exposed to the policy at all and those who had some exposure to the policy ($MD = -.071$, $SE = .024$, $p<.009$), and between those who had some exposure and full exposure to the policy ($MD = -.094$, $SE = .021$, $p<.001$). This means that the percentage of non-ELL students who took at least one AP in the group who had full policy exposure (Group 2) was sixteen percentage points more than the percentage of non-special education students who took at least one AP but had no policy exposure (Group 0). Due to the significant results, the null hypothesis was rejected.

ELL students. There were insufficient numbers of ELL students in the cohorts to determine if there was significance due to the Challenge by Choice Policy.

Overall, the results by subgroup show that some subgroups saw significant increases in AP course taking, e.g., Asian, White, female, male, and Special Education students, which others did not, e.g., Black and Hispanic students.

Table 4

Games-Howell Post-Hoc Test Results by Group

Sample	Group 1	Group 2	*MD*	*SE*	*p*
Overall	None	Some	-0.068*	0.024	0.012
	None	Full	-0.151*	0.021	<.001
	Some	Full	-.084*	0.021	<.001
Asian	None	Some	-0.359*	0.14	0.037
	None	Full	-0.402*	0.124	0.006
	Some	Full	-0.043	0.134	0.945
Black	None	Some	-0.051	0.087	0.832
	None	Full	-0.102	0.075	0.367
	Some	Full	-0.051	0.098	0.861
Hispanic	None	Some	-0.016	0.071	0.972
	None	Full	-0.079	0.06	0.39
	Some	Full	-0.063	0.053	0.463
White	None	Some	-0.066*	0.026	0.032
	None	Full	-0.180*	0.023	<.001
	Some	Full	-0.115*	0.024	<.001
Male	None	Some	-0.073	0.032	0.056
	None	Full	-0.144*	0.027	<.001
	Some	Full	-0.071*	0.028	0.032
Female	None	Some	-0.067	0.035	0.132
	None	Full	-0.17*	0.03	<.001
	Some	Full	-0.103*	0.031	0.003
Special Ed.	None	Some	-0.034	0.024	0.334
	None	Full	-.054*	0.018	0.007
	Some	Full	-0.2	0.03	0.787
Non-Sp. Ed.	None	Some	0.055	0.035	0.075
	None	Full	-.159*	0.022	<.001
	Some	Full	-.104*	0.022	<.001
ELL	None	Some	-0.053	0.053	0.586
	None	Full	0	0	N/A
	Some	Full	0.053	0.053	0.586
Non-ELL	None	Some	-.071*	0.024	0.009
	None	Full	-0.165*	0.021	<.001
	Some	Full	-.094*	0.021	<.001

Note. MD = Mean difference (Group 1- Group 2). *SE* = Standard error of the mean difference.

Table 5 compares the percentages of students taking at least 1 AP course broken down by subgroup. Asian students taking at least 1 AP saw the biggest increase in percentage points (40.2%) followed by White students (18.04%) when comparing those who did not experience the policy at all to the students who had full exposure to the policy. Both Black and Hispanic enrollment increased percentage wise between the pre-policy and post-policy period but did not experience a significant change. Perhaps, the lack of Black and Hispanic enrollment could explain the challenges that these subgroups face even when there are open enrollment Advanced Placement programs. A higher percentage of females (20.3%) when compared to males (17.8%) also enrolled in at least 1 AP class when comparing the groups who had no policy exposure to the group fully exposed. There was also an increase in special education students taking AP classes. No special education students took AP classes during the pre-policy period compared to 5.42% of special education students who enrolled in the post-policy period. That was also reinforced in the ANOVA results which showed a significant difference between special education students who did not experience the policy at all compared to those that were fully exposed.

Table 5

Percentages of Subgroups Taking AP Courses

	None	Some	Full
Asian	30.77%	66.67%	70.97%
Black	6.06%	11.11%	16.22%
Hispanic	21.67%	23.26%	29.58%
White	36.90%	43.48%	54.94%
Special Ed	0%	3.45%	5.42%
ENL	0%	5.26%	0%

Research Question 1B

I also wanted to compare the number of APs students elected to take during high school by group to determine if there were significant gains in the average number of APs students took as a result of the policy. Three groups were again created from the data: no policy exposure (Group 0), some policy exposure (Group 1), and full policy exposure (Group 2).

Again, a one-way between-subjects ANOVA was chosen as the appropriate analysis to determine whether there were any statistically significant differences among the means of the number of APs students elected to take during high school overall and by student subgroup of the three groups. The Levene's Test of Variances was significant for the dependent variable "num_ap" which represented the number of APs students elected to take during high school which indicated that there wasn't homogeneity of

variances, $F(2,3339) = 42.440$, $p < .001$. As such, the Welch's F-test was again used. An alpha level of .05 was used for all subsequent analyses.

The one-way ANOVA of the number of APs students elected to take during high school on the measure of cohort group revealed a statistically significant main effect, Welch's $F(2, 1862.317) = 32.417$, $p < .001$, indicating that not all cohorts had the same average number of APs students elected to take during high school. A Games-Howell post hoc test, as seen in Table 6, also revealed that there was a statistically significant mean difference in the number of APs students took in high school between those who were not exposed to the policy at all and those who had some exposure to the policy ($MD = -.569$, $SE = .150$, $p < .001$), between those that had some exposure to the policy and full exposure to the policy ($MD = -.467$, $SE = .142$, $p = .003$), and between those who had no exposure to the policy and full exposure to the policy ($MD = -1.037$, $SE = .129$, $p < .001$). Therefore, students always exposed to the policy took on average 1 more AP class than those who were never exposed. Due to the significant results, the null hypothesis was rejected.

To answer this question by student subgroup, the researcher first determined if the number of AP courses students decided to enroll in changed because of the Challenge by Choice Policy for males and females. A one- way between-subjects ANOVA was run but the file was split by gender first to determine if there were any statistically significant differences among the means of the number of AP courses males and females enrolled in for the different cohort groups.

Females. The Levene's Test of Variances was significant for the dependent variable "num_ap" which represented the percent of females taking at least one AP

course during high school which indicated that there wasn't homogeneity of variances, $F(2,1572) = 26.825$, $p < .001$.

The one-way ANOVA of the number of AP courses females enrolled in on the measure of cohort group revealed a statistically significant main effect, (Welch's F(2, 906.995) = 25.414, p < .001). Post hoc comparisons, using the Games-Howell post hoc procedure, were conducted to determine which female cohort groups' means differed significantly.

The Games-Howell post hoc test, as seen in Table 6, revealed that there was a statistically significant mean difference in number of classes between the females who were not exposed to the policy at all and those who had some exposure to the policy (*MD* = -.653, *SE* = .216, *p* =.007), between those that had some exposure to the policy and full exposure to the policy (*MD* = -.689, *SE* = .212, *p* =.003), and between those who had no exposure to the policy and full exposure to the policy (*MD* =-1.341, *SE* = .188, *p* <.001). Therefore, the females took on average 1.3 more AP courses in the group who had full policy exposure (Group 2) compared to the females who had no policy exposure (Group 0). Due to the significant results, the null hypothesis was rejected.

Males. The Levene's Test of Variances was significant for the dependent variable "num_ap" which indicated that again, there wasn't homogeneity of variances, $F(2,1764) = 14.841$, $p < .001$. The one-way ANOVA of the percent of males taking at least one AP course on the measure of cohort group revealed a statistically significant main effect, (Welch's F(2, 942.803) = 10.569, $p < .001$).

The Games-Howell post hoc test, as seen in Table 6, revealed that there was a statistically significant mean difference in number of AP classes between the males who

were not exposed to the policy at all and those who had full exposure to the policy (*MD* = -.807, *SE* = .175, *p* >.001). There was also a statistically significant difference between males who experienced no policy and those that experienced some of the policy (*MD* = -.511, *SE* = .206, *p* =.035). Therefore, on average males who were fully exposed to the policy (Group 2) chose to take about .8 more AP classes than the males who had no policy exposure (Group 0). Due to the significant results, the null hypothesis was rejected.

White Students. The Levene's Test of Variances was significant for the dependent variable "num_ap" which indicated that there wasn't homogeneity of variances, $F(2, 2653) = 31.804$, $p<.001$. The one-way ANOVA of the number of AP courses White students took on the measure of cohort group did reveal a statistically significant main effect, Welch's $F(2, 1568.093)= 32.637$, $p < .001$). There was a statistically significant mean difference in the number of AP classes White students took among those who were not exposed to the policy at all and those who had full exposure to the policy (*MD* =-1.171, *SE* = .145, *p* <.001) and also between the Caucasian students who had some exposure to the policy and those that had full exposure to the policy (*MD* =-.604, *SE* = .160, *p* <.001). There was also a significant difference between White students who had some exposure to the policy compared to those that had no exposure to the policy (*MD* =-.5627, *SE* = .164, *p* =.002). Therefore, the White students who had some policy exposure (Group 1) took on average .5 more AP classes than White students who were not exposed to the policy at all (Group 0). Due to the significant results, the null hypothesis was rejected.

Asian Students. The Levene's Test of Variances was not significant for the dependent variable "num_ap" which indicated that there was homogeneity of variances, F(2, 75) = 1.089, *p*=*.342*. The one-way ANOVA of the number of AP courses Asian students took on the measure of cohort group revealed a marginally statistically significant main effect, Welch's F(2, 46.193)= 2.892, *p* =.066). There was a marginally significant difference between Asian students who did not experience the policy at all and those who fully experienced the policy (*MD* =-2.609, *SE* = 1.130, *p* =.063). The results are shown in Table 6.

Black Students. The Levene's Test of Variances was significant for the dependent variable "num_ap" which indicated that there was not homogeneity of variances, F (2, 85) = 4.861, *p*=*.010*. The one-way ANOVA of the number of AP courses Black students took on the measure of cohort group did not reveal a statistically significant main effect, Welch's F(2, 32.263)= 1.941, *p* =.160).

Hispanic Students. The Levene's Test of Variances was significant for the dependent variable "num_ap" which indicated that there was not homogeneity of variances, F (2, 427) = 5.133, *p*=*.006*. The one-way ANOVA of the number of AP courses Hispanic students took on the measure of cohort group did not reveal a statistically significant main effect, Welch's F(2, 145.048)=1.887, *p* =.155).

Non-Special Education Students. The Levene's Test of Variances was significant for the dependent variable "num_ap", which represented the number of AP classes; therefore, there wasn't homogeneity of variances, F(2,3027) = 32.727, $p < .001$. The one-way ANOVA of the number of AP classes non-special education students elected to take on the measure of cohort group revealed a statistically significant main

effect, Welch's F(2, 1697.722) = 32.408, $p < .001$. Post hoc comparisons, using the Games-Howell post hoc procedure, were conducted.

The Games-Howell post hoc test, as seen in Table 6, revealed that there was a statistically significant mean difference in the number of classes between the non-special education students who were not exposed to the policy at all and those who had some exposure to the policy (*MD* = -.533, *SE* = .161, p =.003), between those that had some exposure to the policy and full exposure to the policy (*MD* = -.589, *SE* = .151, p <.001), and between those who had no exposure to the policy and full exposure to the policy (*MD* =-1.122, *SE* = .140, p <.001). Therefore, non-special education students took on average 1.1 more AP classes if they were fully exposed to the policy (Group 2) compared to those who were not exposed to the policy at all (Group 0). Due to the significant results, the null hypothesis was rejected.

Special Education Students. There were not enough Special Education students to run any statistical calculations.

Non-ELL Students. The Levene's Test of Variances was significant for the dependent variable "num_ap", which represented the number of AP classes non-ELL students took during high school based on their cohort, which indicated that there wasn't homogeneity of variances, $F(2,3,248) = 43.021$, $p < .001$. The one-way ANOVA of the number of AP courses non-ELL students took on the measure of cohort group revealed a statistically significant main effect, (Welch's F(2, 1828.718) =35.920, $p < .001$). Post hoc comparisons, using the Games-Howell post hoc procedure, were conducted to determine which non-ELL cohort groups' means differed significantly.

The Games Howell post hoc test, revealed that there was a statistically significant mean difference in number of AP classes non-ELL students took who were not exposed to the policy at all and those who had full exposure to the policy (*MD* = -1.111, *SE* = .131, *p* <.001) between non-ELL students who were not exposed to the policy at all and those who had some exposure to the policy (*MD* = -.595, *SE* = .152, *p*<.001), and between those who had some exposure and full exposure to the policy (*MD* = -.516, *SE* = .145, *p*=.001). This means that non-ELL students who had full exposure to the policy took on average .5 more AP classes than non-ELL students who had some policy exposure. Due to the significant results, the null hypothesis was rejected.

ELL Students. There were insufficient numbers of ELL students in the cohorts to determine if there was significance due to the Challenge by Choice Policy.

Therefore, there was a statistically significant difference between groups in the total number of AP classes students took in high school. For subgroups, the groups differed significantly for females, males, White students, and special education students.

Table 6

Games-Howell Post-Hoc Test Results by Group

Sample	Group 1	Group 2	*MD*	*SE*	*p*
Overall	None	Some	-.569*	0.15	<.001
	None	Full	-1.037*	0.129	<.001
	Some	Full	-0.467*	0.142	0.003
Asian	None	Some	-2.196	1.332	0.238
	None	Full	-2.609	1.13	0.063
	Some	Full	-0.413	1.328	0.948
Black	None	Some	-0.657	0.675	0.603
	None	Full	-0.717	0.407	0.196
	Some	Full	-0.06	0.77	0.997
Hispanic	None	Some	-0.181	0.354	0.866
	None	Full	-0.54	0.304	0.182
	Some	Full	-0.359	0.285	0.421
White	None	Some	-.567*	0.164	0.002
	None	Full	-1.171*	0.145	<.001
	Some	Full	-.604*	0.16	<.001
Male	None	Some	-.511*	0.206	0.035
	None	Full	-.807*	0.175	<.001
	Some	Full	-0.295	0.189	0.261
Female	None	Some	-.653*	0.216	0.007
	None	Full	-1.341*	0.188	<.001
	Some	Full	-.689*	0.212	0.003
Special Ed.	None	Some	-0.121	0.085	0.342
	None	Full	-.133*	0.054	0.039
	Some	Full	-0.012	0.101	0.992
Non-Sped	None	Some	-.533*	0.161	0.003
	None	Full	-1.122*	0.14	<.001
	Some	Full	-0.589	0.151	<.001
ELL	None	Some	-0.053	0.053	0.586
	None	Full	0	0	N/A
	Some	Full	0.053	0.053	0.586
Non-ELL	None	Some	-.595*	0.152	<.001
	None	Full	-1.111	0.131	<.001
	Some	Full	-.516*	0.145	0.001

Research Question 2

My second research question asked: How did the introduction of Challenge by Choice change the trend in AP enrollment and test scores, overall and by student subgroup? To answer the questions about enrollment, two time series analyses was used to analyze changes in (1) the proportion of students taking at least one AP and (2) the average number of AP classes taken by students prior to the policy, at policy implementation, and post-policy implementation. The time series analysis includes three coefficients: the trend in the dependent variable prior to the Challenge by Choice Policy (β_1), the immediate change in dependent variable at the time of policy implementation in 2011 (β_2), and the shift in the trend in the dependent variable post-policy implementation (β_3). To determine the post policy trend, β_2 was added to β_3. Variants of these analyses were also estimated by subgroup where data was sufficient.

To answer the questions about scores, time series models similar to those above were estimated by AP subject with the dependent variable of AP score.

Proportion of Students Taking At Least One AP

Overall, the interrupted time series model explained about 2% of the variance in the percentage of students who took at least one AP exam. This makes sense since the policy change may not have been the only reason the percentage of students took at least one AP class, and many students took APs prior to the onset of the policy. The coefficient for trend prior to policy change (β = -.054, p = .010) suggests that there was a significant decrease in the proportion of students taking APs prior to the policy; in the absence of the policy, this model assumes that downward trend would continue. The coefficient for the

"jump" at the time of policy implementation (β = .165, p < .001) was statistically significant. This shows that the policy change in the 2011 school year led to an average increase of 16.5 percentage points in the proportion of students taking at least one AP course. The coefficient for the shift in the trend post policy change (β = .070, $p \leq .001$) was also significant, suggesting that the percentage of students taking at least one AP began to increase each year post-policy (albeit by a small amount). The results are shown in Table 7.

Female students. The interrupted time series model explained about 2.3% of the variance in the percentage of females who took at least one AP class in high school. The coefficient for the trend prior to policy change (β = -.059, p = .049) suggests that there was a decrease in the proportion of female students taking APs prior to the policy. The coefficient for change at policy implementation (β = .182., p = .010) was statistically significant. This demonstrates that the Challenge by Choice policy led to an increase of 18.2 percentage points in females taking at least one AP course. The coefficient for the trend post policy change (β = .077, p =.012) was also significant, suggesting that the percentage of females taking at least one AP course began to slightly increase in the years after the Challenge by Choice Policy.

Male students. The interrupted time series model explained about 1.8% of the variance in the percentage of males who took at least one AP class in high school. The coefficient for the trend prior to policy change (β=-.045, p=.112) suggests that there was no change in the percentage of males who took one AP. The coefficient for the change at policy implementation (β=.148, p=.024) was statistically significant. This shows that the Challenge by Choice policy led to an increase in 14.8% of males taking at least one AP

course. The coefficient for the trend post policy change (β=.061, *p*=.036) was also significant, suggesting that the percentage of males taking at least one AP course continued to slightly increase after 2011 when the Challenge by Choice policy was implemented. The trends for females and males were very similar however the Challenge by Choice policy led to a greater increase for females than males in taking at least one AP course both at policy implementation and in the post policy years.

White students. The interrupted time series model explained about 2.8% of the variance in the percentage of White students who took at least one AP class in high school. The coefficient for the trend prior to policy change (β=-.055, *p*=.016) was statistically significant. This means that if the policy wasn't put into place the trend would have decreased in the number of White students taking at least one AP course. The coefficient for the trend at policy implementation (β=.160, *p*=.003) was also statistically significant demonstrating that the Challenge by Choice policy led to an increase in 16% of the White students taking an AP course. The coefficient for the trend post policy change (β=.078, *p*=<.001) suggests that the percentage of White students continued to slightly increase after the Challenge by Choice Policy was implemented.

Black students. There was no significance in the model pre-policy, during policy implementation, or post policy. The interrupted time series model explained about 1.9% of the variance in the percentage of black students who took at least one AP class in high school. The coefficient for the trend prior to policy change (β = -.011, p = .871) suggests that there was no change in the proportion of black students taking APs prior to the policy. The coefficient for change at policy implementation (β = .083., p = .611) was also not statistically significant signifying no jump at the time of policy. The coefficient for

the trend post policy change (β = .017, p =.809) was not significant, suggesting that the percentage of black students taking at least one AP course did not change post-policy.

Hispanic students. The interrupted time series model explained about .6% of the variance in the percentage of Hispanic students who took at least one AP class in high school. The coefficient for the trend prior to policy change (β = -.075, p = .288) suggests that there was no change in the proportion of Hispanic students taking APs prior to the policy. The coefficient for change at policy implementation (β = .193., p = .228) was also not statistically significant signifying no jump at the time of policy. The coefficient for the trend post policy change (β = .080, p =.263) was also not significant, suggesting that the percentage of Hispanic students taking at least one AP course did not change post-policy.

The results for the time series analysis for both Black and Hispanic students were similar to that of the ANOVAS showing no significance or change at any time during the policy implementation. This can also show some of the challenges that an open enrollment policy has on subgroups of students. If the policy did not impact subgroups of students, it did not open the door for them.

Non-special education students. The interrupted time series model explained 2.2% of the variance in percentage of non-special education students who took at least one AP in high school. The coefficient for the trend prior to policy change (β=-.059, p=.007) was statistically significant which demonstrates a decreasing trend in the percent of students taking at least one AP prior to the policy. The coefficient for the trend at policy implementation (β=.161, p=.002) was also statistically significant. This shows that the Challenge by Choice policy led to an increase in 16% of the non-special education

students taking an AP course. The coefficient for the trend post policy change (β=.080, p<.001) was again statistically significant trend showing an increase in non-special education students taking an AP course after 2011, when Challenge by Choice was implemented.

Special Education students. The interrupted time series coefficients were not significant for special education students showing that the trend stayed flat before, during, and after the policy.

Non-ELL students. The interrupted time series model explained 2.3% of the variance in percentage of non-ELL education students who took at least one AP in high school. The coefficient for the trend prior to policy change (β=-.055, p=.008) was statistically significant which demonstrates a decreasing trend in the percent of non-ELL students taking at least one AP prior to the policy. The coefficient for the trend at policy implementation (β=.167, p<.001) was also statistically significant. This shows that the Challenge by Choice policy led to an increase in 16.7% of the non-ELL education students taking an AP course. The coefficient for the trend post policy change (β=.075, p<.001) was again statistically significant trend showing a slight increase in non-ELL students taking an AP course after 2011, when Challenge by Choice was implemented.

ELL Students. There were not enough ELL students to analyze the time series analysis results. The results of each subgroup are also shown in Table 7.

Table 7

Summary of Multiple Regression Analysis for Variables Predicting Percent of Students Taking At Least 1 AP Course (N = 3,342)

	Overall	Female	Male	Black	Hispanic	White	Non-ELL	Non-SPED	SPED
Intercept	.237*	.291*	.191*	0.039	0.067	.259*	0.239	0.267	2.69 E-15
Trend Prior	-.054*	-.059*	-0.045	-0.011	-0.075	-.055*	-0.055	-0.059	1.35E-15
Policy	.165*	.182*	.148*	0.083	0.193	.160*	0.167	0.161	0.005
Trend Post	.070*	.077*	.061*	0.017	0.08	.078*	0.075	0.08	0.01

Average Number of APs Taken

Although the percentage of students taking more than one AP was impacted overall by the policy change according to the overall time series analysis, the average number of APs students elected to take did not seem to be impacted by the policy change. The interrupted time series model explained 2% of the variance in the number of APs students elected to take in high school. The coefficient for trend prior to policy change (β=-.010, p=.940) suggests that without the policy change there would be little change in the number of APs students elected to take. The coefficient for change at policy implementation (β=.415, p=.197) was also not statistically significant. This demonstrates that the policy did not change the trend in the number of AP classes students took. The coefficient for the post policy trend (β=.133, p=.346) also showed that there was no change in the number of AP classes students chose to take after the change in policy. All subgroup analyses followed the same pattern as the overall analysis; there were no significant findings. The results are shown in Table 8.

Table 8

Summary of Multiple Regression Analysis for Variables Predicting Number of AP Classes Students Enrolled In (N = 3,342)

	Overall	Females	Males	Black	Hispanic	White	Non-ELL	Non-SPED	SPED
Intercept	1.582*	1.556*	1.627*	0.079	0.2	1.699*	1.599	1.774	3.08E-15
Trend Prior	-0.01	-0.132	0.121	-0.023	-0.35	0.361	-0.014	-0.009	1.69E-15
Policy	0.415	0.678	0.161	0.798	0.88	-0.001	0.417	0.341	0.083
Trend Post	0.133	0.307	-0.044	0.007	0.411	0.156	0.152	0.161	0.01

Note: $*p < .05$.

AP Scores by Subject

The scores for each subject were then analyzed and the results are in Table 9. A brief review of the findings is provided here.

Biology/Living Environment. The interrupted time series model explained 9.7% of the variance in the scores of biology students. The trend in scores was flat prior to the policy (β = -.296, p = .093). At the onset of the policy, scores increased on average by about 1 point (β =1.062, p = .015) and scores began to increase slightly post policy (β = .368, p = .039). This is important to note because although the mean percentage of students (10.3%) who took biology from 2011 to 2016 increased from the mean percentage of students who took biology beforehand (6.8%), the scores also went up at the time of the policy and afterward.

Calculus AB. The interrupted time series model explained 3.0 % of the variance in scores for Calculus AB students. Before the policy, scores were decreasing slightly (β=-.406, p=.005). The coefficient for change at policy implementation (β=.630, p=.064) shows a flat trend or no significant jump at the time the policy went into effect. The

coefficient for the post policy trend however (β=.370, p=.010) shows that scores jumped about .38 points after the policy went into effect. Again, importantly this demonstrates that although more students entered into a Calculus AB class overall, the scores did not decrease. Before 2011, the mean percentage of students who took Calculus AB was 12.8% compared to 2011 to 2016 (policy implementation and post policy period), when the mean was 13.5% of students. In fact, the trend moved from a decreasing trend before policy toward a flat trend post policy in scores.

Calculus BC. The scores for Calculus BC did not show any significance at any point including before, during, or after the policy change. Before the policy, the mean percentage of students taking Calculus BC was 5.8% compared to during the policy and after when the mean percentage was 9%. Therefore, more students entering the class during the policy and post policy did not change the scores significantly.

English. A similar result occurred for English scores as Calculus BC. There was not any significance at any point that the policy had on English scores. There was also no significance before or after the policy. However, there was on average 4.2% more students who took the AP English Language and Composition after the policy was implemented demonstrating that more students taking the class did not cause a decrease in the trend of test scores.

Macroeconomics. Macroeconomics, however, produced statistically significant results. The interrupted time series model explained 2.1% of the variance in scores for Macroeconomics students. The coefficient for trend prior to policy change (β=.438, p<.001) shows a significant positive trend in scores before the policy. The coefficient for change at policy implementation (β=-1.239, p<.001) shows that scores decreased on

average by about 1.2 points. The coefficient for the post policy trend (β=-.440, p<.001) shows that scores dropped on average by about .44 points after the policy went into effect which really showed that the trend stayed flat after decreasing at the time of policy. Although there is a decreasing trend at the time of the policy, macroeconomics also saw the largest increase on average (15.175%) in the percentage of students who took the class from the time Challenge by Choice was implemented into the post policy time period. Such a large increase in the number of students makes sense that the scores might decrease especially without a covariate for prior performance. These results are shown in Table 9.

Table 9

Summary of Multiple Regression Analysis for Variables Predicting AP Scores (N = 3,342)

	Living Env.	English Lang.	U.S. Hist.	World Hist.	Macro	Calc AB	Calc BC
Intercept	1.506*	3.176*	3.169*	3.566*	3.851*	2.340*	3.278*
Trend Prior	-0.296	0.079	0.173	0.117	.438*	-.406*	0.163
Policy	1.062*	0.008	-0.511	-0.132	-1.239*	0.63	-0.051
Trend Post	.368*	-1.443	-0.179	-0.171	-.440*	.379*	-0.176

Note: *$p < .05$.

Overall, AP scores for six out of the seven subjects studied either increased on average at the time of policy implementation or stayed flat. The fact that none of the scores decreased in these classes even though more students started enrolling in them shows that heterogenous mixing of students did not cause a decline in scores. The only course that saw a decline in scores at the time of policy implementation was Macroeconomics. Similarly, in post policy years, most subjects studied, except for Macroeconomics, either saw a significant positive increase in scores or score trends remained flat. Macroeconomics was the only subject that saw a significant decrease in scores, however it was also the subject that had the largest increase in enrollment in post-policy years.

CHAPTER 5

Implications of Findings

Prior to the Challenge by Choice Policy, students only accessed AP classes by passing certain prerequisites such as a GPA in a prior class and teacher recommendation. The Marxist view suggests that this policy, by closing off access to certain students, reinforces the social inequalities in society during students' education. And the findings from this research support that claim. Prior to the policy fewer students took AP courses compared to post-policy, suggesting that there were students who wanted to take the courses but were unable to. The policy effect was quite large: Students who started high school after the policy was in place were 18.6 percentage points more likely to take at least 1 AP class than students who graduated prior to the policy. This suggests a large utilitarian benefit of the policy.

However, when the data for this study was broken down by race, the results indicate that the Challenge by Choice policy functioned to increase AP take-up only among some student groups. Because the ANOVA was higher powered than the time series analysis, the researcher used those results to determine the significance of the policy for each ethnicity group. White students who had full exposure to the policy were more likely to enroll in APs, by 21.4 percentage points, than those who had no exposure to the policy. However, for Black and Hispanic students, the Challenge by Choice Policy did not have any significant impact on the percentage taking at least 1 AP course according to the ANOVA results. Notably, however, the output might lack significance for Black students because of the small sample size. There was also no policy effect for the ELL or Special Education populations. This suggests that while the policy did

effectively increase AP course-taking in a utilitarian sense, not all groups benefitted equally from it. In other words, there is room for improvement in regard to equity.

There was mixed evidence as to whether the policy led to students taking more AP courses. The ANOVA results suggest that students are taking more APs; however, this was not corroborated in the time series analyses perhaps due to small annual samples. The ANOVA suggested that students who were fully exposed to the policy took an average one more AP class than those who were never exposed to the Challenge by Choice policy. But, again, there were differences among subgroups: White, Asian, and Special Education students took significantly more AP courses after the policy, while there was no significant increase in the number of APs for Hispanic or Black students. It is also important to note again that the sample size for the minority populations was quite small. Larger samples of students in each subgroup would be needed to determine any impact that this policy might have had on each group.

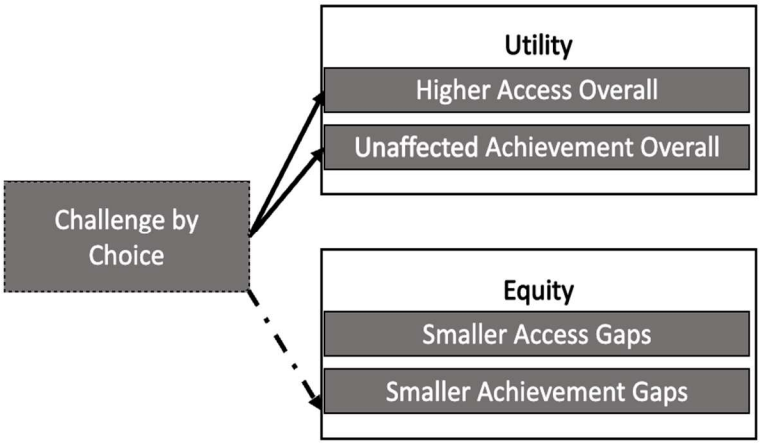
Finally, despite an increase in the numbers of students – and arguably an increase in the number of students who were less academically prepared – taking AP courses, there was not a significant drop in AP scores, except in Macroeconomics. This suggests that the policy increased access without sacrificing performance – contrary to the argument many make in favor of tracking. The fact that achievement was not impacted even though more students were given access and opportunity to more challenging classes supports the Marxist philosophy of education and Rawls's Theory of Justice.

Overall, the policy was effective in a utilitarian sense—more students took AP courses post-policy. However, there is still question of whether the policy can be effective in an equity sense, as not all subgroups appeared to benefit or benefit equally.

More may need to be done to ensure that students from all demographic backgrounds are encouraged to take APs.

The study's conceptual model was corroborated by some of the results; according to the results, Challenge by Choice is essentially a policy that led to higher access overall to challenging classes and did not impact achievement. Challenge by Choice, a social reconstructionist idea, opened access for more students who would have been previously barred from taking Advanced Placement classes. However, the results also showed that Challenge by Choice was actually a model of higher utility that did not yet meet the equity model. Since I did not have the power to determine if achievement gaps for subgroups were impacted, the dot in the conceptual model remains dotted or unknown. The results of the study are shown in the model in Figure 2.

Figure 2
Conceptual Model Results



Therefore, Challenge by Choice was a more utilitarian policy in that it benefitted the student body overall. However it did not close access gaps in terms of equity and achievement gaps would have to be further researched.

Comparison to Prior Literature

The findings herein are similar to other studies which show detracking efforts that open access to all students can increase overall enrollment in AP classes (Burris & Garrity, 2008). However, the findings also align with research that argues that even in open enrollment or choice programs, minority students still might not take advantage and enroll in the more challenging classes. Mickelson and Everett (2008) found that open enrollment occupational tracks reproduced the social stratification in society; variations in enrollments reflected the race, ethnic, and social class of the students in North Carolina. Other studies found that open enrollment programs still produced equity gaps because previous tracking efforts created cultures among the students that could not be broken just by a new policy (Yonezawa, Wells, & Serna, 2002). In another study, Lofton found high African enrollment in a sixth-grade class because it was completely detracked. However, seventh and eighth grade math and science classes were self-selected and were comprised of all White students (Lofton, 2019). The study refers to this idea as the "duplicity of equality" because while a school might seem to offer more equitable programs, it might not close the gap in minority enrollment (Lofton, 2019, p. 25).

The results show that there are still equity gaps that the Challenge by Choice policy did not impact. In fact, the gap widened between White and Black students by 7.8 percentage points and between White and Hispanic students by 10.13 percentage points when comparing the groups that had no policy exposure to those that had full policy exposure. In the class of 2019, 63% of White students took at least 1 AP but only 27% and 16% of Hispanic and Black students took 1 AP respectively. In the same class, 0%

of ENL students and 14% of special education students took an AP class. The equity gaps in the school are reflective of the disparities in AP/IB enrollment overall in New York State. In schools statewide, Black, Latinx and low-income students are underenrolled; 27% of low-income students are enrolled while 55% of not low-income students are enrolled in AP/IB. Furthermore, 27% of Black students and 28% of Latinx students are enrolled while 46% of White students are enrolled in AP/IB courses ("Course Access: A Statewide Challenge," n.d.). As the school looks to close equity gaps, we can learn from studies that show that there has to be a further push for students to take certain higher-level classes to close equity gaps (Lofton, Yonezawa, Wells, & Serna, Michelson & Everett, Burris & Garrity, 2008). This finding underscores Rawls' Theory of Justice Lens, which suggests that subgroups must be given more weight for a policy to produce fair results.

The findings that scores did not change with increased access further aligns with previous research by Atteberry, Lacour, Burris, Welner, and Murphy (2019) who found that detracking was not associated with lower mean math IB performance in any group even though there was increased participation. It also supports Vygotsky's Zone of Proximal Development and constructivist theory. Vygotsky believed that the social aspects of learning and thus, the heterogeneous grouping of all achievement levels helps all students.

Limitations of the Study

With an interrupted time series analysis, one of the limitations is that other factors were occurring at the same time as the policy change and could have led to some of the changes. For example, changing demographics, new technology in the classrooms,

changes in administration and teaching staff, etc. could have led to changes in enrollment and achievement of students in AP classrooms unrelated to the policy itself. In terms of curriculum, universal acceleration for Algebra occurred in 8th grade in 2007 and for Earth Science in 2011. Although this did not impact the high school specifically, it did impact course taking behavior at the high school. The universal acceleration policy is not an opt in policy like Challenge by Choice. Every eighth-grade student, because of universal acceleration, began taking both Algebra and Earth Science leaving students more of a chance to take both AP Science and AP Math courses at high school. There were also turnovers in staffing, specifically in some of the teachers who taught the courses that were reviewed in this study. Moreover, the school demographics have changed over the years. The number of Hispanic students and the ENL population have increased greatly since before the policy. In 2005, 6% of the school was Hispanic and 3% were ENL students. In 2019, 20% of the school was Hispanic and 4.6% were ENL students. This is important context in understanding the results of the policy change.

Additionally, this study did not factor in prior achievement or socioeconomic status as variables in the analysis. If there was a covariate in the study based on prior achievement, the researcher would have been able to determine how the policy impacted low, middle, and high achieving students. This is important, particularly because one of the arguments against detracking is that it negatively impacts high achieving students. Therefore, it would be important for the researcher to determine how the Challenge by Choice policy specifically impacted the achievement of high achieving students. The researcher would also want to ascertain how the policy impacted low and middle

achieving students in enrolling in an AP class and their subsequent achievement in a more challenging class.

Also, the researcher only used one suburban district in New York State and so the findings are limited to other small, suburban districts in New York State with similar demographics. The sample for this study was largely white (*n*=2,653) and Hispanic (*n*= 427) and the small sample size of Black (*n*=87), Asian (*n*=77), ELL (*n*=90), and Special Education (*n*=311) students limited the statistical power of the analyses. Limited sample size may lead to type two errors which sometimes cause a null hypothesis to be retained when it should not be. This may explain the differences between the ANOVA analyses, which pooled pre and post samples, and time series analyses, which used annual samples.

Recommendations for Future Practice

Importantly, school leaders and other staff members must consistently look at enrollment and achievement data in classes through both a utilitarian lens and a Theory of Justice lens. Open enrollment programs are one solution to opening access to advanced coursework for all students. But it still does not answer the problem of increasing minority enrollment in AP classes if the students do not opt to enroll themselves. Findings from this study show that after Challenge by Choice, AP participation increased but the participation rate for minority students is still not equal. Scores also did not change significantly except for one class upon policy implementation. One approach that the school can take is to auto enroll students into classes who meet benchmark proficiency levels on state examinations, such as the Regents, or in prerequisite courses. Students who meet proficiency would automatically be enrolled in the subsequent highest available class including AP classes. Therefore, instead of relying on students enrolling

themselves, automatic enrollment removes the barriers that keep students out of advanced placement coursework. Parents can still opt their students out of the advanced coursework but the prevailing notion is that all students should be achieving at high levels in advanced coursework. Those systemic and institutional barriers would, however, be removed.

The new automatic enrollment policy would eliminate all barriers that bar students out of AP courses. One high school in Spokane, Washington started to change their approach in 2008 and adopted a district wide acceleration policy in 2015. By 2015, all students who showed mastery on their state reading, science, and math tests were automatically placed in the next available rigorous course. About half the students in the high school came from low-income homes and the school hoped to increase enrollment in their AP courses by adopting this policy of auto-enrollment. Other districts to implement an automatic enrollment policy were the Federal Way and Tacoma districts. In 2009-10, 41 percent of all of Federal Way's 11th and 12th graders took an advanced course and 35 percent of minority students enrolled. However, the year after the district adopted an automatic acceleration policy, the numbers rose to 67 percent for all students and 61 percent for minority students (Blad, 2020). Washington state fought to pass a law that requires students who meet or exceed certain grade level standards or exams be placed into the next most rigorous advanced course offered (Ed Trust, 2019). The law was passed and schools had to adopt an academic acceleration policy by the 2021-22 school year. North Carolina and Colorado similarly have passed such policies in recent years so that the traditional underrepresented students are automatically placed into rigorous courses and do not have to enroll themselves (Ed Trust, 2019). This ensures all students

who were “unidentified” previously as AP students take courses that challenge them (Blad, 2020).

In addition, automatic enrollment cannot be the only support that students and staff receive along the way to remove the opportunity gaps. Vertical alignment of curriculum, professional development to help teachers reach students of all abilities in their classrooms, and mentorship for both staff and students are important to ensure the automatic enrollment placement is successful. Reviewing curriculum at each grade level by administrators and teachers ensures that students are consistently getting instruction that builds on concepts from the grade prior and that they are prepared for advanced level coursework in high school. School leaders and teachers must be constantly monitoring curriculum auditing as well as the enrollment and achievement data of students to determine the success of such modifications. School leaders should also regularly ask for feedback from teachers, students, and families to determine the successes and challenges of growing AP programs.

Recommendations for Future Research

One possible area for future research might be completing a comparative study between similar schools who implement a tracking policy, an open enrollment policy, and an auto enrollment policy. By conducting a comparative study, the number of external factors occurring at the same time can be controlled. If the researcher controls for as much as possible in each setting, then the three systems can be compared and the policy’s impact on enrollment and achievement can be studied. Thus, the researcher might gain a deeper understanding of how each policy impacts the overall student body and individual subgroups.

Another possible avenue for future research might be to conduct a study on how prior achievement and socioeconomic status impacts student enrollment and achievement in AP courses after an open enrollment policy is implemented. Therefore, the impact of the policy on high, middle, and low achieving students would be portrayed as well as students from differing socioeconomic backgrounds. Importantly, this research could show how an open enrollment program impacts the test scores of high achieving students. Even though this study showed that the scores did not decrease, except in macroeconomics, it did not show what happened specifically to students who scored highly before the policy went into effect. Being that a main argument to continue tracking is that detracking harms high achieving students, the community would want to know if high achieving students are impacted by the policy change. Also, it might tell us the success rate of students who were not high achievers in AP classes and how to best support them. It would also show how many students, who would have been previously barred out of AP classes, were now able to access them.

Lastly, a qualitative analysis of teacher and student perspectives of the Challenge by Choice policy might give further insight into the programs' successes and challenges. A case study of this district would give further detail on why students might enroll in AP classes while others face barriers. Staff and student perspectives on the policy's effectiveness might also provide important information moving forward for the district. If there is not a sense of community embracing and supporting the policy change, it would be difficult for it to be successful. A case study might also give insight into the challenges or successes teachers have faced in the classroom with students of varying achievement levels. This information could help administrators be aware of areas of need

and potential ways to address challenges that the staff and students perceive are results of the policy.

Conclusion

This study showed that moving to a Challenge by Choice open enrollment policy for AP classes improved the enrollment of students taking at least one AP course and to some extent the average number of AP courses students chose to take overall. In other words, it opened access to students in the high school in a utilitarian sense, with minimal effects on achievement. However, from an equity standpoint, it did not have as much of an impact across racial groups: Hispanic, Black and ELL students were not impacted after the policy went into effect. As demographics change, so does the responsibility of schools to ensure that every student receives equal opportunities. The onus is on schools to "level the playing field" so that the hierarchy of society is not simply reinforced by the system. This means that detracking must start early on in the schooling process. Policy changes such as Challenge by Choice cannot begin sporadically, but rather should be a gradual shift within the context of the school's vision focused on increasing equity for subgroups of students.

Overall, an open enrollment program for AP classes, such as Challenge by Choice, is a first step for a district to take a step in providing open access to challenging classes regardless of prior achievement, race, socioeconomic status, gender, etc. However, more is needed—just because students can enroll in more challenging classes doesn't mean they always will. Therefore, school leaders must be aware of this pitfall and ensure that supports are in place for the students and staff and that students are encouraged and/or automatically placed in challenging classes. Building a community

vision increasing both enrollment and equity must be constantly revisited to ensure that a program such a Challenge by Choice is really benefiting all students, particularly those often underrepresented.

REFERENCES

A Guide for Parents of English Language Learners in New York State. (n.d.). Retrieved March, 2021, from http://www.nysed.gov/common/nysed/files/programs/bilingual-ed/guideforparentsofellsinnysenglish.pdf

About AP Scores. (n.d.). Retrieved March, 2021, from https://apstudents.collegeboard.org/about-ap-scores

Alvarez, D., & Mehan, H. (2006). *Whole-school detracking: A strategy for equity and excellence*. San Diego: University of California.

Anderson, S., Medrich, E., & Fowler, D. (2007). Which Achievement Gap? *Phi Delta Kappan*, *88*(7), 547–550. https://doi.org/10.1177/003172170708800716.

AP Data - AP Report to the Nation 2014 – Research – College Board. (2018, February 16). Retrieved July 07, 2020, from https://research.collegeboard.org/programs/ap/data/nation/2014.

AP Data – Archived Data 2008– Research – College Board. (2018, February 21). Retrieved July 07, 2020, from https://research.collegeboard.org/programs/ap/data/archived/2008.

Argys, L., Rees, D., & Brewer, D. (1996). Detracking America's Schools: Equity at Zero Cost? *Journal of Policy Analysis and Management, 15*(4), 623-645. Retrieved April 14, 2021, from http://www.jstor.org/stable/3326053.

Atteberry, A., Lacour, S. E., Burris, C., Welner, K., & Murphy, J. (2019). Opening the Gates: Detracking and the International Baccalaureate. *Teacher College Record*, 121, 1-63.

Barnard-Brak, L., McGaha-Garnett, V., & Burley, H. (2011). Advanced Placement Course Enrollment and School-Level Characteristics. *NASSP Bulletin*, 95(3), 165-174. doi:10.1177/0192636511418640.

Blad, E. (2020, December 8). The simple policy change that's getting more students of color in advanced courses. *Education Week*. Retrieved March 9, 2022, from https://www.edweek.org/leadership/the-simple-policy-change-thats-getting-more-students-of-color-in-advanced-courses/2020/03.

Boaler, J. (2011). Changing Students' Lives Through the Detracking of Urban Mathematics Classrooms. *Journal of Urban Mathematics Education, 4*(1), 7-13.

Boaler, J., Wiliam, D., & Brown, M. (2000). Students Experiences of Ability Grouping - disaffection, polarisation and the construction of failure. *British Educational Research Journal, 26*(5), 631-648. doi:10.1080/713651583.

Bowles, S., & Gintis, H. (1976). *Schooling in capitalist America*. London.

Bowles, S., & Gintis, H. (2001, November 08). *Schooling in Capitalist America Revisited*. Retrieved February, 2021, from https://www.umass.edu/preferen/gintis/soced.pdf.

Bridgeman, B., Morgan, R., & Wang, M. (1996). Reliability Of Advanced Placement Examinations. *ETS Research Report Series, 1996*(1), I-9. doi:10.1002/j.23338504.1996.tb01681.

Burdman, P. (2000). Extra credit, extra criticism. *Black Issues in Higher Education.* 17 (8), 28-33.

Chiang, I. A., Jhangiani, R. S., & Price, P. C. (2015). *Research Methods of Psychology*. Retrieved from https://opentextbc.ca/researchmethods/front-matter/about-this-book/.

Clark, D. (2010). Selective Schools and Academic Achievement. *The B.E. Journal of Economic Analysis & Policy, 10*(1). doi:10.2202/1935-1682.1917.

College Board. (2016a). AP program participation and performance data. Retrieved from https://research.collegeboard.org/programs/ap/data/participation/ap-2016.

Conant, J. B. (1961). *Slums and Suburbs. A commentary on schools in metropolitan areas.(*Fourth printing.). New York: McGraw-Hill Book.

Course access: A statewide challenge. The New York Equity Coalition. (n.d.). Retrieved February 5, 2022, from https://equityinedny.edtrust.org/course-access-a-statewide-challenge/.

Counts, G. S. (1932). *Dare the school build a new social order?* Carbondale: Southern Illinois University Press.

Counts, G. S. (1928). *School and Society in Chicago.* New York, NY: Harcourt, Brace, and Co.

Counts G. S. (1922). T*he Selective Character of American Secondary Education*. Chicago: University of Chicago Press.

Counts, G. S. (1926). *The Senior High School Curriculum*. Chicago: University of Chicago Press.

Counts, G. S. (1927). *The Social Composition of Boards of Education: A Study in the Social Control of Public Education.* Chicago: University of Chicago Press.

Creswell, J. W., & Guetterman, T. C. (2019). *Educational research: planning, conducting, and evaluating quantitative and qualitative research*. Harlow, England: Pearson.

Curry W, MacDonald W, Morgan R. The Advanced Placement program: access to excellence. *Journal of Secondary Gifted Education*. 1999;11(1):17-23. doi:10.4219/jsge-1999-612.

Dewey, J. (1958). *Democracy and Education*. New York, NY: The Macmillan Company.

Dewey, J. (1937). *Education and Social Change*. Bulletin of the American Association of University Professors, 23(6), 472-474.

Dewey, J. (1897). *My pedagogic creed and the demands of sociology upon pedagogy* by prof. Albion W. Small (with an introd. by S.T. Dutton). Chicago: A. Flanagan.

Dewey, J. (1899). *The School and Society*. Chicago, IL: University of Chicago Press.

Domina, T., McEachin, A., Hanselman, P., Agarwal, P., Hwang, N., & Lewis, R. (2016). Beyond Tracking and Detracking: The Dimensions of Organizational Differentiation in Schools. doi:10.7249/wr1155.

Dougherty, C., Mellor, L., & Jian, S. (2006). The relationship between advanced placement and college graduation. Austin, TX: *National Center for Educational Achievement.*

Duflo, E., Dupas, P., & Kremer, M. (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *The American Economic Review*, 101(5), 1739-1774. http://dx.doi.org.jerome.stjohns.edu:81/10.1257/aer.101.5.1739.

Ed Trust. (2019, December 9). *5 things to advance equity in access to and success in advanced coursework*. The Education Trust. Retrieved March 9, 2022, from https://edtrust.org/resource/5-things-to-advance-equity-in-access-to-and-success-in-advanced-coursework/.

Figlio, D. N. & Page, M. E. (2002). School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality?, *Journal of Urban Economics, 51*(3), 497-514.

Garrity, D. T., & Burris, C. C. (2008). *Detracking for Excellence and Equity*. Association for Supervision and Curriculum Development.

Hargrove, L., Godin, D., & Dodd, B. (2008). College outcomes comparisons by AP® and non-AP high school experiences (Research Report No. 2008-3). Austin, TX: *College Board*.

Hebel, S. (1999). AP courses are new target in struggle over access to college in California. *Chronicle of Higher Education,* 46(14), A32-A34.

Kelly, S. (2007). The contours of tracking in North Carolina. *High School Journal*, 90, 15-31.

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.

Khine, K. (2019, May 10). *A Greater Number of Jobs Require More Education, Leaving Middle-Skill Workers with Fewer Opportunities.* Retrieved from http://statchatva.org/2019/05/10/a-greater-number-of-jobs-require-more-education-leaving-middle-skill-workers-with-fewer-opportunities/.

Kolluri, S. (2018). Advanced Placement: The Dual Challenge of Equal Access and Effectiveness. *Review of Educational Research, 88*(5), 671-711. doi:10.3102/0034654318787268 .

Kulik, J. A., & Kulik, C. C. (1992). Meta-analytic findings on grouping programs. *Gifted Child Quarterly*, 36, 73-77. doi:10.1177/001698629203600204.

Lavy, V., Paserman, M. D., & Schlosser, A. (2008). Inside the Black of Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom. *NBER Working Papers 14415, National Bureau of Economic Research*. doi:10.3386/w14415.

Lefgren, L. (2004). Educational peer effects and the Chicago public schools. *Journal of Urban Economics, 56*(2), 169-191. doi:10.1016/j.jue.2004.03.010.

Lewis, T., & Cheng, S. (2006). Tracking, Expectations, and the Transformation of Vocational Education. *American Journal of Education, 113*(1), 67-99. doi:10.1086/506494.

Lofton, Jr., R. (2019). The Duplicity of Equality: An analysis of Academic Placement in a Racially Diverse School and Black Community. *Teachers College Record, 121*, 1-48.

Loveless, T. (2016). How Well Are American Students Learning? *Brown Center on Education Policy at Brookings, 3*(5), 1-40.

Mattern, K., Marini, J., & Shaw, E. (2013). Are AP® Students More Likely to Graduate from College on Time? *College Board,* 1-31. Retrieved October 8, 2020, from https://files.eric.ed.gov/fulltext/ED556464.pdf.

Mattern, K. D., Shaw, E. J., & Xiong, X. (2009). The relationship between AP Exam performance and college outcomes (College Board Research Report No. 2009-4). New York, NY: *College Board.*

McBride Davis, C. A., Slate, J. R., Moore, G. W., & Barnes, W. (2015). Advanced placement exams, Incentive Programs, and cost effectiveness: A lack of equity and excellence for black students in Texas, New York, and Florida. *The Journal of Negro Education*, *84*(2), 139. https://doi.org/10.7709/jnegroeducation.84.2.0139.

Mickelson, R. A., & Everett, B. J. (2008). Neotracking in North Carolina: How High School Courses of Study Reproduce Race and Class-Based Stratification. *Teachers College Record, 110*(3), 535-570.

Mill, J. S. (1863). *Utilitarianism*. London: Parker, Son and Bourn.

More colleges stop giving credit for AP exams. (2019, April 27). Retrieved October 9,2020, from https://www.marketplace.org/2013/02/20/education/more-colleges-stop-giving-credit-ap-exams/.

More Schools Offering Open Enrollment for AP Courses. (2013, November 04). Retrieved February, 2021, from https://rossieronline.usc.edu/more-schools-offering-open-enrollment-for-ap-courses/.

Murphy, D., & Dodd, B. (2009). A Comparison of College Performance of Matched AP® and Non-AP Student Groups. Research Report No. 2009-6. *College Board,* 1-52. Retrieved October 8, 2020, from https://files.eric.ed.gov/fulltext/ED561022.pdf.

Oakes, J., & Guiton, G. (1995). Matchmaking: The dynamics of high school tracking decisions. *American Educational Research Journal, 32*, 3–33.

Oakes, J., Wells, A. S., Jones, M., & Datnow, A. (1997). Detracking: The social construction of ability, cultural politics, and resistance to reform. *Teachers College Record, 98*(3), 482–510.

Opportunity Gap Definition. (2013, September 03). *The Glossary of Education Reform*, Retrieved from https://www.edglossary.org/opportunity-gap/.

Patrick, K., Sokol, A., & Morgan, I. (2020). Inequities in advanced coursework: What's driving them and what leaders can do. *Education Trust*, 1250 H Street NW Suite 700, Washington, DC 20005. Retrieved from Social Science Premium Collection.

Promotion of advanced placement (2000). *Gifted Child Today Magazine.* 23(3), 6.

Rawls, J. A. (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Riehl. C, Pallas, A. M. & Nariello, G. (1999). Rites and Wrongs: Institutional explanations for the student course-scheduling process in urban high schools. *American Journal of Education*, 107(2), 116-154. Retrieved from: http://www.jstor.org/stable/1085482.

Rosenbaum, J. E. (2000). If Tracking Is Bad, Is Detracking Better? *American Educator,* 1-7.

Rubin, B. C. (2006). Tracking and Detracking: Debates, Evidence, and Best Practices for a Heterogeneous World. *Theory Into Practice, 45*(1), 4-14. doi:10.1207/s15430421tip4501_2.

Santoli, S. P. (2002). Is There an Advanced Placement Advantage? *American Secondary Education, 30*(3), 23-35.

Schneider, J. (2009). Privilege, equity, and the Advanced Placement Program: Tug of war. *Journal of Curriculum Studies*, 41(6), 813-831. doi:10.1080/00220270802713613.

Spade, J. Z., Columba, L., & Vanfossen, B. E. (1997). Tracking in mathematics and science: Courses and course-selection procedures. *Sociology of Education*, 70, 108-127.

Student Enrollment - AP Central: College Board. (2021, March 02). Retrieved March, 2021, from https://apcentral.collegeboard.org/exam-administration-ordering-scores/ordering-fees/manage-registration-ordering/access-setup-ap-registration-ordering/student-enrollment.

Theokas C. & Saaris, R. (2013). Finding America's Missing AP and IB Students. *The Education Trust*. Washington, D.C.

Tracking and Ability Grouping in Middle Level and High Schools. (2020, June 04). Retrieved February, 2021, from https://www.nassp.org/tracking-and-ability-grouping-in-middle-level-and-high-schools/.

Vygotskij L.S., & Cole, M. (1981). *Mind in society: the development of higher psychological processes*. Cambridge, Mass. Harvard Univ. Press.

Warne, R. T. (2017). Research on the Academic Benefits of the Advanced Placement Program: Taking Stock and Looking Forward. *SAGE Open*. https://doi.org/10.1177/2158244016682996.

Whiting, G. W., & Ford, D. Y. (2009). Multicultural Issues: Black Students and Advanced Placement Classes: Summary, Concerns, and

Recommendations. *Gifted Child Today*, 32(1), 23–26. https://doi.org/10.4219/gct-2009-840.

Wiatrowski, M., Hansell, S., Massey, C., & Wilson, D. (1982). Curriculum Tracking and Delinquency. *American Sociological Review, 47*(1), 151-160. doi:10.2307/2095049

Yonezawa, S., & Jones, M. (2006). Students Perspectives on Tracking and Detracking. *Theory Into Practice, 45*(1), 15-23. doi:10.1207/s15430421tip4501_3.

Yonezawa, S., Wells, A. S., & Serna, I. (2002). Choosing tracks: "Freedom of choice" in detracking schools. *American Educational Research Journal*, 39, 37–67.

Zarate, Maria & Pachon, Harry (2006). Equity in offering advanced placement courses in California high schools, 1997-2003 Gaining or Losing Ground? New York: Tomas Riveria Policy Institute.

Vita

Name	*Jessica Williams*
Baccalaureate Degree	*Bachelor of Arts*, *Hofstra University, Hempstead* *Major: Social Studies Education/History*
Date Graduated	*May, 2008*
Other Degrees and Certificates	*Advanced Certificate in Educational Leadership (2014),* *Master of Liberal Studies, Stony Brook University, Stony Brook,* *Major: Liberal Studies*
Date Graduated	*August, 2010*