Magnet Programs, "Mission Creep," And Mitigating Opportunity Gaps:

An Analysis of The Relationship Between Magnet Program Themes and Racial Achievement Gaps

by
Madeleine Lewis Peery

A thesis submitted to the Department of Education Leadership and Policy Studies,
College of Education
in partial fulfillment of the requirements for the degree of

Executive Ed.D.

in K-12 Professional Leadership

Chair of Committee: Dr. Bradley Davis

Committee Member: Dr. Virginia Rangel

Committee Member: Dr. Catherine Horn

Committee Member: Dr. Geovanny Ponce

University of Houston
August 2021

Dedication

I dedicate this thesis to my students - past, present, and future. Over the past three years, my driving purpose in this program has been learning how to shape schools and school districts to provide the educational opportunities you need to grow and thrive. I promise to continue to push myself to learn, reflect, and improve so I can be the best school or district leader I can be for you.

Further, I dedicate this work to my mother, whose love taught me how to serve the people around me and whose passing taught me to cherish every minute with them. I could not have accomplished this degree without you.

Acknowledgments

First and foremost, I wish to thank Dr. Davis for his guidance throughout my doctoral journey, particularly with regards to this thesis. Your feedback, encouragement, methodological expertise, and sage practicality were integral to my completion of this work. Thank you in particular for recognizing my passion and potential and setting high expectations for this thesis. Undeniably, you have made me a better scholar and empowered me as a practitioner.

I would like to thank my other three invaluable committee members - Dr. Rangel, Dr. Horn, and Dr. Ponce. Your recommendations provided highly beneficial perspectives that aided me in improving the quality of this thesis and its significance within my field.

This doctoral journey would not have been possible without my eight supportive, insightful, dedicated cohort-mates. To the future Drs. Coello, Constant-Garza, Hill-Kennedy, Perry, Riordan, J. Tovar, M. Tovar, and Vaughan, I extend my greatest thanks for your contributions to my education and for your friendship. I look forward to the countless years of learning and laughing together that lie ahead of us.

Additionally, I wish to thank my co-workers at Shadowbriar Elementary School, Durham Elementary School, and J.R. Harris Elementary School. Your support and flexibility as I navigated working full time while completing my master's and doctoral programs have meant so much to me.

Lastly, I am beyond grateful for my siblings, Caroline and Henry; my mother-in-law, Lori; and my husband, Travis. You have been my constant cheerleaders, reminding me that I was capable of accomplishing this feat when it felt most daunting. Travis, thank you for creating time for me to read and write by taking on the bulk of the responsibilities

at home, as well as for reminding me to take a break once in a while. I love you all more than I can express, and I could not have achieved this dream without you.

Abstract

Background: In 2017, nearly 3.5 million students attended a magnet school (Magnet Schools of America, 2017) - a public school offering a specialized curriculum to diverse groups of students. Magnets were born from the school desegregation movement of the 1960s, but, within and beyond the context of magnet programming, racial diversity in schools has been increasingly deprioritized since the mid-1970s. This raises particular concern as racial achievement gaps continue to pervade American public education, fueled by gaps in the disparate learning opportunities afforded to White students and students of color. Often, magnet programming is promoted as a potential remedy to opportunity and achievement gaps, yet this aspiration has yet to be sufficiently explored, much less substantiated. **Purpose:** The purpose of this work was to explore the racial achievement gaps within magnet schools compared to nonmagnet schools. Further, I aimed to fill a significant gap in the literature by identifying which - if any - magnet themes (e.g., STEM, fine arts, Montessori) contribute to or mitigate within-school racial achievement gaps. I engaged in this work in hopes of developing actionable recommendations for school districts regarding equitable magnet programming that minimizes racial opportunity gaps. **Methodology:** In this quantitative study, I leveraged multiple regression to analyze the relationships between magnet programming and within-school racial achievement gaps in the 11 “major urban” public school districts in Texas. The campus-level sample consisted of 1,357 schools, representing nearly 950,000 students. I aggregated demographic and STAAR testing data published by the Texas Education Agency and analyzed district and campus websites to determine which schools offer magnet programs and with what themes. I calculated within-school racial achievement gaps in math and reading standardized test scores at the fifth grade and end-of-course (high school) levels. Using these data, I conducted multiple regressions to distill the relationships between magnet programs, as well as magnet themes, and within-school racial achievement gaps. **Results:** Within Texas's 11 major urban school districts, certain magnet themes appeared to contribute to within-school racial achievement gaps while others appear to lessen them by as much as 48.2 percentage points. In particular, elementary STEM/STEAM, elementary leadership, and secondary early college magnets predicted decreases in multiple within-school gaps, while elementary fine arts magnets and secondary language magnets related to an increase in such disparities. 10 of the 17 magnet themes in the sample were not found to significantly predict an achievement gap in any model analyzed. **Conclusion:** This work indicated that magnet programming has influenced within-school racial achievement gaps at urban public schools in Texas. Importantly, my analysis has addressed a significant gap in the literature by identifying specific magnet themes that mitigate and exacerbate these gaps, offering direct implications to school and district leaders in terms of equitable magnet programming.

Keywords: magnet schools, magnet programs, magnet themes, achievement gaps, opportunity gaps

Table of Contents

Chapter Page

List of Tables

Table Page

List of Figures

Figure Page

Chapter I

Introduction

The story of magnet programs is a complex one, fraught with noble intentions, political implications, competing priorities, and mixed efficacy. The U.S. Department of Education defines a magnet school as a public school “that offers a special curriculum capable of attracting substantial numbers of students of different racial backgrounds” (U.S. Department of Education, 2017, para. 2). Magnets vary substantially in terms of their school level (e.g., elementary, secondary), format (whole-school vs. program-within-school), application policies, and curricular or pedagogical theme. Despite this heterogeneity, one element tends to remain constant: controversy. Magnet programs “serve as microcosms of many competing educational policy values in the United States today; values of choice, high standards, equity, diversity, and desegregation all compete simultaneously at the doorsteps of magnet schools” (Goldring, 2009, p. 362). Considering the intersection of these complex issues, the lack of consensus on the effects of magnet programs, and - above all - the nearly 3.5 million students who attend a magnet school each year in the United States (Magnet Schools of America, 2017), magnet programs remain a vitally important research topic today.

Background

Magnet programs initially emerged in the late 1960s and 1970s in response to federal courts ordering public school districts to racially desegregate their schools (Blank et al., 1983, 1996; Blank & Archbald, 1992; Clotfelter, 2004; Frankenberg & Le, 2008; Frankenberg & Siegel-Hawley, 2010; Poppell & Hague, 2001; Rossell, 2005; West, 1994). Districts that had been ordered to dismantle de jure segregation policies recognized magnet programs as palatable integrators with the potential to improve racial balance while minimizing White flight (Blank et al., 1996; Blank & Archbald, 1992; Clotfelter, 2004; Frankenberg & Le, 2008; Frankenberg & Siegel-Hawley, 2010; West, 1994). In addition to judicial pressures, federal grants, including the Emergency School Aid Act and later the Magnet Schools Assistance Program, incentivized desegregation and magnet program implementation (Blank et al., 1983; Christenson et al., 2003; Clinchy, 1993; Clotfelter, 2004; Frankenberg, 2008; Frankenberg et al., 2008; Frankenberg & Le, 2008; Goldring, 2009; Magnet Schools Assistance Program, 1985; Orfield & Lee, 2007). The degree to which magnet programs have impacted school segregation remains unclear though, with a history of contradictory findings in the literature, detailed in Chapter Two of this thesis.

However, contemporary magnet programs serve distinctly different and broader purposes than those of the 1970s and 1980s. Significant "mission creep" (Frankenberg & Le, 2008) has occurred, such that today's magnet policies and programs tend to prioritize academic performance and curricular enrichment over desegregation (Fleming, 2012; Frankenberg et al., 2008; Frankenberg & Le, 2008; Frankenberg & Siegel-Hawley, 2010; Goldring & Smrekar, 2000; Orfield & Lee, 2007; Rhea & Regan, 2007; Rossell, 2003; Siegel-Hawley & Frankenberg, 2012, 2013; Smrekar & Goldring, 2011; Wang et al., 2014; Yang et al., 2005). This change has been catalyzed by a number of factors including competition among school choice options (Frankenberg et al., 2008; Frankenberg & Siegel-Hawley, 2010; Goldring, 2009; Wang et al., 2014), high-stakes testing (Fleming, 2012; Frankenberg et al., 2008; Goldring, 2009; Rossell, 2003; Smrekar & Goldring, 2011), and the deprioritization of school desegregation by federal courts and legislation (Betts et al., 2015; Fleming, 2012; Frankenberg et al., 2008; Frankenberg & Le, 2008; Hannah-Jones, 2014; Orfield & Lee, 2007; Rossell, 2003; Siegel-Hawley & Frankenberg, 2012; Smrekar & Goldring, 2011). As with their efficacy as tools for desegregation, no consensus exists regarding magnet programs' impact on students' academic opportunities and outcomes, except that "in at least some times and places, students benefit from enrolling in magnet schools" (Ballou, 2007, p. 31). I explore the heterogeneity of these findings in Chapter Two as well.

Nonetheless, magnet programs currently represent the largest sector of school choice in the United States, enrolling more students than charter schools, private schools, homeschooling, or voucher programs (Frankenberg et al., 2008; Goldring, 2009; Siegel-Hawley & Frankenberg, 2012, 2013). In fact, close to 3.5 million students attended magnet programs in 2017, indicating that around one in 15 American public school students are enrolled in a magnet (Magnet Schools of America, 2017). Yet, despite their prevalence, much remains uncertain about magnet programs - their efficacy as a desegregation strategy, their impact on academic opportunities and achievement, and their role within the future of American education. Further exploration of these issues is highly warranted, given the millions of students attending magnet programs each year. Knowing only that "in at least some times and places, students benefit" (Ballou, 2007, p. 31) is simply insufficient, particularly when much is on the line for students, particularly the students whom school desegregation - arguably the reason magnet programs exist - was intended to affect most.

Problem Statement

Students of color in the United States have historically been subjected to and continue to be subjected to objectively inferior academic opportunities than their White peers (Alemán, 2006; Aud et al., 2011; Berends, 2014; Darling-Hammond, 1997; Gilliam, 2016; Gilliam et al., 2016; Gregory et al., 2006; Haycock, 2001; Kahlenberg, 2001; Ladson-Billings, 2010; Lee, 2002; Manning & Kovach, 2002; Mattison & Aber, 2007; Mickelson, 2001; Milner, 2010; Noguera & Wing, 2006; Oakes, 1986; Orfield et al., 2008; Orfield & Lee, 2007; Petras et al., 2011; Rothstein, 2004; Rubin et al., 2006; Scherff & Piazza, 2009; Skiba et al., 2002, 2011; Swanson, 2004; U.S. Department of Education, Office of Civil Rights, 2014; Wang, 1998; Williams, 2002a, 2002b; Wing, 2006; Wu et al., 1982). Researchers have identified grave, long-term ramifications of this systemic disparity, including significantly reduced grades (Mickelson, 2001), standardized test scores (Aud et al., 2011; Mickelson, 2001; Orfield & Lee, 2007), high school graduation rates (Aud et al., 2011; Swanson, 2004), college enrollment rates (Aud et al., 2011), and educational attainment among students of color (Mickelson, 2001). Considered together, the systematic withholding of educational opportunity and the consequences of this inequity create a feedback loop that structurally oppresses students of color.

Researchers and practitioners alike have explored innumerable reforms to potentially address opportunity gaps, including magnet programs. However, despite their potential for school diversification and academic enrichment, the degree to which magnet programs mitigate such gaps remains unclear. Magnet research is rife with limitations including research age, single-district samples, inconsistent variable operationalization, poorly controlled variables, possible confounding variables, and misaligned dependent variables (Ballou et al., 2006; Ballou, 2009; Betts et al., 2015; Blank & Archbald, 1992;

Harris, 2006; Reardon & Owens, 2014; Rothstein, 2004; Siegel-Hawley & Frankenberg, 2013; Wang et al., 2014; Wang & Herman, 2017; Yang et al., 2005). Additionally, magnet research tends to compare academic outcomes at magnet schools to comparable nonmagnet schools, but no research appears to have compared racial disparities in student outcomes within magnet schools. Further, researchers have called for exploration of possible differences in outcomes across magnet themes, a significant gap with major implications for schools and districts (Betts et al., 2015; Bifulco et al., 2009; Gamoran, 1996; Wang et al., 2014, 2017; Wang & Herman, 2017). One can imagine that STEM, fine arts, International Baccalaureate, career preparation, and foreign language programs - to name a few - do not offer identical educational opportunities resulting in equivalent academic outcomes for students. Unfortunately, the extant research continues a historical trend of merely "describing the average of this diverse category of schools" (Gamoran, 1996, p. 14). Such questions regarding magnet programs endure, and, while they remain unanswered, school districts and campuses must individually attempt to determine what policies and practices will best serve students.

Given the perpetuation and severe consequences of the racial opportunity gap and the subsequent disparities in academic outcomes, it is imperative that school districts and state education agencies implement research-supported interventions to disrupt the toxic cycles of the status quo. Williams (2002) noted that "many educators continue to seek the single program … that will improve all students' achievement" and mitigate this gap (p. 19). Although this panacean program has yet to (and may never be) identified, magnets are at times marketed as a possible remedy for the opportunity gap. This theory persists despite a lack of research exploring the effects of magnet programs on gaps in educational opportunity and academic outcomes. Without further understanding of the effects of magnet programs, including the benefits or limitations of specific themes, the education community will continue to shoot in the dark - a dangerous practice where students' lives are involved.

Purpose Statement

The purpose of this nonexperimental multiple regression analysis is to first determine if there is a relationship between magnet status (whether or not a campus offers a magnet program) and within-school racial achievement gaps when controlling for other relevant variables known to impact academic outcomes. Further, I will examine whether this relationship differs between magnet program themes, holding numerous campus-level variables constant, filling the problematic gap in the literature described above. These findings may shed light on how school districts should develop and implement educational offerings to their students in an effort to close racial opportunity gaps.

Research Questions

In this study, I seek to answer the following research questions:

Research question 1: What is the nature of the relationship between magnet status (i.e., whether a school offers a magnet program or not), controlling for contextual variables known to influence academic outcomes?

Research question 2: Does this relationship vary across magnet program themes?

Through the first research question, I aim to determine whether or not schools with magnet programs have measurably different disparities in state standardized test scores compared to schools without magnet programs, controlling for relevant contextual variables in an effort to distill the possible effects of the magnet program itself on student outcomes. The second research question addresses the wide variety in magnet themes by disaggregating the work of the second research question by magnet theme. Considered together, I intend to leverage these two questions to contribute to the collective knowledge surrounding a larger question on the minds of researchers and practitioners alike: do magnet programs mitigate opportunity gaps?

Significance of the Work

In an interview with Fleming (2012), magnet researcher Claire Smrekar identified the twenty-first century as “a pivotal time for school districts and education leaders to clearly define the role of magnet schools” (p. 1). Magnet programs are the largest sector of school choice in the country, as measured by student enrollment (Frankenberg et al., 2008; Goldring, 2009; Siegel-Hawley & Frankenberg, 2012, 2013), and it’s likely that they will continue to impact millions of students each year. Additionally, with each year that the opportunity gap persists, more and more students fall through the cracks, and, as stated in the majority ruling of *Brown v. Board* (1954), “it is doubtful that any child may reasonably be expected to succeed in life if he is denied the opportunity of an education” (p. 493). Magnet programs are often marketed as a potential tool to mitigate this gap, leading to better academic outcomes for students of color. However, this has yet to be conclusively established, nor has the extant research determined which of the many magnet themes make such an impact. Without these determinations, school and district decision makers can only use anecdotal findings, tangentially relevant research, and guesswork to develop impactful magnet programs and policies. Magnet programs and opportunity gaps touch the lives of millions of children each year, therefore guidance must be available for districts and campuses as soon as possible regarding how magnet programs of various themes can mitigate or magnify educational disparities.

Overview of Framework

The opportunity-to-learn (OTL) framework shifts perspective from academic outputs to educational inputs (Dowd et al., 2014; Griffin et al., 2007; Guiton & Oakes, 1995). This framework argues that “academic performance, no matter the students' backgrounds, can change with an improvement in the quality of instruction” (Stevens, 1993, p. 4). Although researchers have evaluated different variables that impact students’ opportunities to learn, the four primary pillars of the OTL framework are content coverage, content exposure, content emphasis, and quality of instructional delivery (Stevens, 1993, 1996; Wang, 1998). By addressing these modifiable inputs as opposed to “achievement gaps,” policymakers and practitioners can shift their focus to methods for improving educational practice in service of students.

The OTL perspective is highly relevant to research examining magnet programs and their effects on opportunity gaps, as magnet programs are a frequently-implemented form of modifiable inputs. I utilize this lens as I measure the relationship between magnet programs - an academic input - and racial disparities in standardized test scores - an academic outcome - to deliberately maintain focus on the opportunities school and district decision makers enable for their students, especially their students of color. Addressing research questions similar to those of this thesis using an achievement gap mentality simply misses the point and risks blaming students for gaps in achievement. According to Milner (2010), “focusing on an achievement gap inherently forces us to compare culturally diverse students with White students without always understanding reasons that undergird disparities” (p. 8). In contrast, using the OTL framework “helps us to move the discourse from the exhausted 'achievement gap' lament to a more robust and nuanced discussion of why school failure persists for some groups of students" (Ladson-Billings, 2010, pp. ix-x). I seek to clarify whether magnet programs help to break the cycle of inequitable educational opportunity, a calling which directly aligns with the OTL framework.

Overview of Methodology

To address the research questions listed above, I take a quantitative approach, analyzing data collected and published by the Texas Education Agency. This publicly-available database shares key data from each public school district and campus within the state, including student demographic data and state standardized test scores, disaggregated by student demographic groups. The majority of the nation's magnet programs are offered in urban communities (Wang & Herman, 2017), so, to focus on urban magnet program implementation, I analyze data from the eleven districts labeled as "major urban" districts by the Texas Education Agency due to their locations and student enrollment totals. In the 2019-20 school year, these districts ranged in enrollment from over 40,000 students to over 200,000, representing 51 to 280 campuses each. In total, I analyze campus-level data from 1,360 schools serving nearly 950,000 students. I code the schools as magnet or nonmagnet based on magnet program lists published on each district's school choice website, a resource I use to code magnet schools by their respective themes as well.

After downloading the relevant data and coding the 1,360 campuses by magnet status and theme, I run statistical tests using Stata. To answer my first research question, I conduct a multiple regression analysis to determine the relationship between a campus's status as magnet or nonmagnet and the disparity between the standardized test scores of its students of color and its White students, controlling for variables known to impact student achievement. To explore the second research question, I adapt the model used to answer the first question, replacing the binary magnet/nonmagnet independent variable with a disaggregated magnet theme variable. The details of this plan - including information regarding the sample, variable operationalization, control variables, regression assumptions and procedures, are provided in Chapter Three.

Limitations

The greatest limitation within the extant magnet research is the inherent misalignment between the key inputs of magnet programming and the primary dependent variables studied, and it is worthwhile to examine that concern within the context of this thesis. Virtually all magnet studies measure magnet programs' effects in terms of standardized test scores and/or graduation rates, given the practical challenges of measuring how a magnet theme enriches learning experiences at a given campus. However, "many magnet school programs are not specifically aimed at building the skills reflected on standardized tests" (Blank & Archbald, 1992). A well-developed, thoroughly-implemented fine arts magnet, for example, may be unlikely to significantly affect students' scores on multiple-choice reading and math tests. This discrepancy between the focus of magnet programs and the focus of most magnet research introduces an inherent validity concern. However, the research questions addressed in this thesis do in fact align with standardized test scores, mitigating this validity concern to some degree. My curiosity lies with the degree to which magnet programming impacts math and reading standardized test scores, as these data influence what doors are open to students as they leave K-12 education.

In Chapter Five, I will expound upon these limitations as well as others that impact this work. The data presently available do not enable a causality attribution, as a great deal of heterogeneity exists across magnet programs in ways that cannot easily be quantified, publicized, aggregated, and controlled for within statistical analyses. For example, within the context of this research design, I am unable to explore the impact of the inevitable variations of the depth and frequency of magnet programming across magnets. I also do not have access to the individual application and selection policies each campus adopts, which influences which students receive magnet learning opportunities. These considerations should be kept in mind while interpreting the results of this thesis, which aims to illuminate one facet of how magnet programs manifest academic opportunity and enrichment. Again, I address limitations in Chapter Five.

Outline of Remaining Chapters

In the following chapters, I will explore the extant literature regarding magnet programs, then set and follow a methodological pathway to explore my research questions. Chapter Two reviews current and historical literature regarding the history of magnet programs and their documented effects on school- and student-level outcomes, as well as research regarding the opportunity gap and the associated opportunity-to-learn framework. In Chapter Three, I delve into the details of my methodology, including justifying the methods selected. Chapter Four offers an analysis of the data, and Chapter Five consists of conclusions and implications based on the results.

Chapter II

Review of the Literature

As stated in Chapter One, magnet programs "serve as microcosms of many competing educational policy values in the United States today; values of choice, high standards, equity, diversity, and desegregation all compete simultaneously at the doorsteps of magnet schools" (Goldring, 2009, p. 362). One cannot examine the effects of magnet programs on students' learning opportunities without the background of the interplay between these various elements. To provide such a perspective, in Chapter Two, I offer a brief overview of the relevant literature, spanning the history, current reality, educational opportunities, and academic effects of magnet programs. Each element represents a cornerstone of the foundation necessary for a contextualized analysis of magnet programs today.

In the first portion of this chapter, I summarize the history of public school desegregation in the United States, in terms of both case law and federal funding, as magnet programs would likely not exist without this turning point in American public education. This paradigm shift catalyzed the conception of magnet programs and shaped the culture and values present in early magnet development. When political forces reversed course by minimizing and even repealing mandates for school desegregation, the purpose of magnet programs shifted correspondingly. I also explore the history of funding for magnet programs, including how federal grants have followed the trajectory of case law by subsidizing and then deemphasizing integrated learning environments. With this context established, I then outline the extant research on how magnet programs have helped or hindered school desegregation over the years. Understanding the noble roots of magnet programs - enabling school desegregation - and how that effort was systematically deconstructed over the years is a necessary foundation for those striving to gauge the value of magnet programs today.

With this history in mind, I then turn to contemporary magnet programs. I first share insights from magnet researchers regarding the current mission of magnets, including the consensus that this mission has changed over time. Specifically, that mission has "crept" from addressing opportunity gaps to emphasizing academic enrichment, typically without consideration of students' backgrounds. I then outline information on contemporary magnet applications, waitlists, and demand. Next, the research on school choice includes analyses of how and why families choose schools. The data reveal that families consider the race and ethnicity of students already enrolled in schools, not merely schools' academic quality, furthering a fundamental precept of this thesis established in the first section - issues of race and equity underlie the past and present of American public education, including magnet programming. Considered together, the contemporary mission of magnets, their application processes and demand for enrollment, and families' school choice approaches illustrate the relationship between magnet programs and their clientele today, as well as how racial equity concerns impact and are impacted by that relationship.

Next, I explicitly address the opportunity gaps present and persistent in public education in the United States. I explore measures, sources, and outcomes of such gaps in this section. The pervasiveness and ramifications of opportunity gaps drives my use of the opportunity-to-learn (OTL) theoretical framework for this research. The OTL framework centers researchers' and practitioners' perspectives on how modifiable academic inputs directly impact students' learning opportunities. This enables an evaluative frame of mind that encourages us to reflect on what we can do, as those empowered to shape education, to foster equitable learning opportunities for all students. With this perspective, I undertake the research at hand - not merely for the sake of learning about magnets but also to reveal actionable avenues to improve magnet programming to benefit historically underserved students.

With these three crucial elements established - the history of magnets, their present reality, and the context of opportunity gaps - I then present the documented effects of magnet programs on student achievement. These findings carry more weight when considering the background of America's tumultuous school desegregation past, the current relationship between magnets and their stakeholders, and the disparities in educational opportunity across the nation's public schools. To further contextualize these results, I also describe school-level factors that can impact magnet programming. Finally, I identify a major gap within the extant literature - a lack of research exploring the potentially differential impacts of magnet programs of various themes on achievement gaps. Armed with an understanding of magnets' past, their present, and the high stakes of mitigating opportunity gaps, I aim to address that gap.

Racial Desegregation: A Federal Priority

The history of magnet programs is situated within the larger history of racial desegregation in United States public schools, and we must look to the past to understand the present. Over two crucial decades, the federal government engaged in a "root-to-branch excavation of racial discrimination" (Deo, 2008, p. 444) as a number of Supreme Court cases, legislative policies, and funding initiatives compelled school districts to dismantle school segregation. One avenue to comply with this federal mandate was magnet programs, which were supported politically and financially by the federal government. Since then, magnets have become the largest sector of school choice in the United States in terms of student enrollment (Frankenberg et al., 2008; Goldring, 2009; Siegel-Hawley & Frankenberg, 2012, 2013).

Case Law and Legislation from 1954 to 1973

The story of American school desegregation begins with the landmark case *Brown v. Board of Education of Topeka, Kansas I* (1954), in which the U.S. Supreme Court unanimously struck down *Plessy v. Ferguson*'s (1896) "separate but equal" doctrine, which had permitted de jure segregation. At the time, de jure - or "by law" - segregation policies were present in schools throughout the South, the Border states (Kentucky, Missouri, West Virginia, Maryland, and Delaware), the District of Columbia, and select districts in other regions of the country (Clotfelter, 2004; Deo, 2008; Frankenberg, 2008; Frankenberg & Siegel-Hawley, 2010; Holley-Walker, 2010; Orfield et al., 2014). The following year, in *Brown v. Board of Education of Topeka, Kansas II* (1955), the Court unanimously ruled that states and school districts must comply with *Brown I* "with all deliberate speed" (p. 301).

However, few Southern districts actively pursued desegregation following these two cases, let alone with any speed. Virtually all participated in a "Massive Resistance" by merely removing race-based enrollment restrictions and offering voluntary school choice programs (Clotfelter, 2004; Frankenberg et al., 2008; Frankenberg & Siegel-Hawley, 2010; Welner, 2006). These policies enabled Black and White families to transfer to historically all-White or all-Black schools, respectively, of their own volition.

Few families chose to exercise these options, resulting in stagnant racial segregation throughout the late 1950s and the 1960s (Clotfelter, 2004; Frankenberg, 2008; Johnson, 2011; Lutz, 2011). In 1963, nine years after *Brown I*, 99% of Black students in the South still attended segregated schools (Orfield et al., 2014). This resistance was primarily limited to the South, as most districts in the Border states dismantled their de jure segregation policies promptly (Clotfelter, 2004).

President Lyndon B. Johnson, a Southerner himself, countered the Massive Resistance by spearheading the 1964 Civil Rights Act (Orfield et al., 2014; Orfield & Lee, 2007). Title VI of the Act tied federal funding to desegregation measures through its nondiscrimination mandate, while Title IV authorized the federal government to sue districts refusing to desegregate (The Civil Rights Act of 1964). The primary source of this federal funding was the Elementary and Secondary Education Act (ESEA, 1965). Effectively, the Civil Rights Act and the ESEA offered education funding as a “carrot” while threatening litigation and removal of that financial support as a “stick.” This accountability system enabled greater enforcement of civil rights law “than any administration before or since” (Orfield et al., 2014)

In 1968, the U.S. Supreme Court wielded the “stick” authorized in the Civil Rights Act through *Green v. County School Board of New Kent County* (1968). In its unanimous decision in *Green*, the Court ordered districts to take active steps to address de jure segregation. This could be achieved through mandatory student assignment policies, including the use of busing, but the Justices determined passive practices such as New Kent County, Virginia’s voluntary school choice plan to be inadequate. *Green* also established a six-part operational definition of sufficient desegregation for districts to be granted “unitary status”: student bodies, faculties, staffs, facilities, transportation, and extracurricular activities that are no longer identifiable by race. With this ruling, the Supreme Court effectively forced active compliance with *Brown I*, passed 14 years prior. For good measure, one year after *Green*, the Court ruled in *Alexander v. Holmes County Board of Education* (1969) that school desegregation must occur immediately in the South, another unanimous decision in response to Massive Resistance techniques.

Two major cases in the early 1970s expanded the scope of school desegregation. In yet another unanimous decision, the U.S. Supreme Court reaffirmed the use of mandatory student assignment and busing as a desegregation method in *Swann v. Charlotte-Mecklenburg Board of Education* (1971). The Justices also encouraged the use of numerical goals for student and faculty desegregation in this ruling. However, the Justices provided states and districts with a degree of discretion in this process and implied that such measures were to be temporary. Two years later, in *Keyes v. Denver School District No. 1* (1973), the Court extended desegregation goals to districts with de facto segregation - segregation that occurs without official government or school district authorization. *Keyes* also expanded the reach of desegregation goals beyond addressing Black-White separation in the South, as this case involved Latino students in Colorado. Taken together, these cases represented a judicial push towards actively addressing racial segregation in all of its forms throughout the country’s public schools.

Court-Ordered Desegregation Plans

Credit is due to private civil rights organizations, including the NAACP, for initiating lawsuits against school districts that engaged in racial discrimination. To build legal precedent while minimizing public backlash, these organizations targeted “easier wins" first, such as admitting individual Black students to graduate programs, and later pursued systemwide, paradigm-shifting cases such as *Brown I* (Clotfelter, 2004; Guryan, 2004; Johnson, 2011). This helped catalyze the government's use of lawsuits for desegregation accountability. Many cases, particularly those against school districts, resulted in court orders to desegregate. This practice is still in use today, albeit with less frequency and intensity, a fact I address below in the section "A Wave of Unitary Status Declarations."

Primarily initiated by the federal government (Clotfelter, 2004), these court-ordered desegregation plans typically assign active desegregation techniques - including busing, magnet programs, and joint city-suburban approaches - for segregated school districts (Betts et al., 2006; Frankenberg & Lee, 2002). These plans "often [entail] extensive micromanagement, very often extending down to the detail of specifying precise attendance boundaries" (Clotfelter et al., 2006, p. 367). Further, most desegregation orders include a specific numerical target for racial desegregation (Rossell, 2003). These orders reinforce that federal funding is dependent upon desegregation, pursuant to Title VI of the Civil Rights Act (Hannah-Jones, 2014; The Civil Rights Act of 1964). Prior to the 1964 Civil Rights and the 1965 ESEA, on average, districts did not comply with court-ordered desegregation until a decade after receiving the order, consistent with the obduracy of the Massive Resistance (Johnson, 2011). As of 1965, however, the average time between receiving a court order and implementing its plan decreased to less than two years (Johnson, 2011). From this point on, court-ordered desegregation plans began to effect change.

Effects of Court-Ordered Desegregation Plans. Through large-scale quantitative analyses, researchers have demonstrated that court-ordered plans have led to a decrease in racial segregation in schools (Clotfelter, 2004; Reardon & Owens, 2014). In fact, court orders have been the greatest driver of desegregation in U.S. education history (Reardon & Owens, 2014). Court-ordered desegregation has had a ripple effect on districts not under order as well, such that "the shift in the legal and social environment and enforcement by political leaders contributed to declining segregation in nearly all districts" (Reardon & Owens, 2014, p. 206) in the South.

The desegregation catalyzed by these court orders has been remarkably impactful for students of color. Johnson (2011) compared a number of long-term outcomes among 9,000 Black and White individuals who were of school age before, during, and after desegregation court orders in their districts. Johnson found that Black individuals who were young when their district implemented a court-ordered desegregation plan (thus experiencing desegregated education for most of their school years) had substantially greater educational attainment than Black individuals who experienced more years of segregated schooling. In addition, Johnson demonstrated that court-ordered desegregation increased Black students' college quality, occupational prestige, adult earnings, adult socioeconomic status, and adult health, as well as decreasing incarceration rates (Johnson, 2011). No differences were found among White students, contrary to desegregation opponents' fears that White students would "suffer" from interactions with Black peers, nor were differences found within districts that did not adopt a court-ordered desegregation plan (Johnson, 2011). Taken together, these data reveal that compliance with court orders to racially desegregate can make a significant, lasting difference for Black students.

The Emergency School Aid Act (ESAA)

In 1972, the Emergency School Aid Act was implemented to financially support districts that needed additional resources to develop and enact desegregation plans (Blank et al., 1983; Christenson et al., 2003; Clotfelter, 2004; Frankenberg et al., 2008; Frankenberg & Le, 2008; Orfield & Lee, 2007). Only districts with plans to reduce racial isolation among students by at least five percent were eligible for ESAA funds, which could be used on professional development (particularly for teachers who had only experienced one-race schooling), curriculum development, and school desegregation research (Blank et al., 1983; Frankenberg, 2008). Beginning in the 1970s, many districts used ESAA grants to adopt an increasingly popular mechanism for enabling school-level diversity: magnet programs (Blank et al., 1996; Christenson et al., 2003; Frankenberg, 2008).

Implementation of Magnet Programs

Palatable Integrators. When confronted with mandates to end racial segregation, a number of districts implemented magnet programs as they “offered a relatively uncontroversial - and peaceful - means of integrating schools” (Rossell, 2005, p. 46). Magnet programs became particularly popular in large Southern school districts, where White families “defecting to private, parochial, or suburban school districts” (Poppell & Hague, 2001, p. 3) impeded compliance with desegregation mandates. In effect, magnets served as a compromise, offering expanded and novel academic programming to attract students to help diversify schools’ student bodies (Blank et al., 1996; Blank & Archbald, 1992; Frankenberg & Le, 2008; Frankenberg & Siegel-Hawley, 2010; West, 1994). Districts aimed to primarily incentivize enrollment among White families, which some regarded as “the magnet consumers to be pleased” (West, 1994, p. 2574). Metropolitan districts primarily sited magnet programs at campuses historically attended by students of color (Clotfelter, 2004), which White families were unlikely to attend without the lure of increased academic opportunity. In this way, magnet programs offered the potential to overcome challenging barriers to racial desegregation in schools, including White flight.

Early Magnet Programs. Early magnet programs were modeled after “select” public schools located in large cities, including the Boston Latin School, the Bronx School of Science, and Lowell High School in San Francisco (Blank et al., 1983, 1996; Blank & Archbald, 1992; Doyle & Levine, 1984; Frankenberg & Le, 2008; West, 1994). The first magnet program on record was established in 1968 at McCarver Elementary, located in Tacoma, Washington (Goldring, 2009; Kafer, 2012; Rossell, 2005). Prior to the implementation of its magnet program, McCarver Elementary’s student population was 91% Black, despite the fact that only 4% of Tacoma residents were Black (Rossell, 2005). The district engaged in a door-to-door publicity campaign about the magnet program and recruited top educators to staff it, leading to an immediate decrease in racial isolation in its first year with the magnet program, when minority enrollment dropped to 64% (Rossell, 2005). The following year, the nation’s second magnet program opened in Boston (Kafer, 2012; Rossell, 2005), and, 13 years later, over 1,000 magnet programs existed across the nation (Blank et al., 1983).

Magnet Programs Popularized. In 1977, only 14 districts offered magnet programs (Blank et al., 1983). A mere five years later, in 1982, the number had risen to over 1,000 magnet programs serving 441,000 students (Blank et al., 1983, 1996; Smrekar & Goldring, 2000). As of 1992, over 3,100 magnet programs enrolled over 1.2 million students, with a majority of the nation's suburban and urban school districts offering magnet programs (Blank et al., 1996; Blank & Archbald, 1992; Smrekar & Goldring, 2000; Steele & Eaton, 1996). By 2001, the total number of magnet programs exceeded 5,500 (Christenson et al., 2003). The most recent reliable national data indicate that, as of 2017, nearly 3.5 million students attend one of the country's over 4,300 magnet programs (Magnet Schools of America, 2017). In comparison to charter schools, private schools, voucher programs, and homeschooling, magnet programs now represent the largest sector of school choice in the United States (Frankenberg et al., 2008; Goldring, 2009; Siegel-Hawley & Frankenberg, 2012, 2013).

Racial Desegregation Deprioritized

Beginning in the 1970s, the tides turned in the realm of public school desegregation "from a root-to-branch excavation of racial discrimination into a benign pruning around the edges" (Deo, 2008, p. 444). New case law, changes in funding, and a wave of unitary status declarations removed the teeth from earlier desegregation policies, altering the manner in which magnet programs could impact students. This shift in priorities continues to impact magnet students' learning opportunities today.

Case Law from 1974 to 2007

The U.S. judicial system's pattern of active support of school racial desegregation was broken in 1974. That year marked a drastic turning point due in large part to the four new U.S. Supreme Court Justices, including the Chief Justice, appointed by President Richard Nixon (Clotfelter, 2004). In a 5-4 decision (with Nixon's recent appointees representing four of the five majority votes), *Milliken v. Bradley* (1974) barred a mandatory interdistrict student assignment plan in the heavily segregated Detroit metropolitan statistical area. In Detroit, White families had fled to suburban districts, leaving Black students isolated downtown. Although district boundaries directly contributed to de facto segregation, the Justices ruled that interdistrict student assignment plans could only be employed if suburban districts directly contributed to de jure segregation. This ruling “sealed off the boundaries between many American cities and their suburbs, creating an easy (and nearby) alternative for white parents fleeing desegregation orders in urban centers” (Frankenberg et al., 2008, p. 10). Following this ruling, racially distinct urban and suburban districts could no longer collaborate in desegregation efforts, enabling a “return to the doctrine of ‘separate but equal’ for urban school children” (Orfield & Lee, 2007, p. 8).

Following *Milliken*, magnet programs became increasingly prevalent. Federal courts rapidly approved their use as desegregation tools, beginning with *Morgan v. Kerrigan* (1976). A number of similar cases reaffirmed the use of magnet programs and voluntary majority-to-minority transfers, in which students could transfer to any school in which their race represents the minority (*Amos v. Board of School Directors of City of Milwaukee*, 1976; *Armstrong v. Board of School Directors of the City of Milwaukee*, 1976; *Arthur v. Nyquist*, 1976; *Ross v. Houston Independent School District*, 1983). Voluntary practices, such as the implementation of magnet programs, partially filled the vacuum of mandatory desegregation practices left in the wake of *Milliken*.

However, the greatest blow to school desegregation was a trifecta of U.S. Supreme Court cases in the early 1990s which drastically lowered the bar for districts under desegregation court orders to be declared unitary and removed from court supervision (Clotfelter et al., 2006; Frankenberg, 2008; Holley-Walker, 2010;

Kahlenberg, 2011; Lutz, 2011; Orfield & Lee, 2007; Reardon et al., 2012; Reardon & Rhodes, 2011; Reardon & Yun, 2003). These three rulings "collectively express[ed] the opinion that the courts [had] 'done enough' in the area of school desegregation" (Lutz, 2011, p. 134). First, the Court ruled in a 5-3 decision in *Board of Education of Oklahoma City v. Dowell* (1991) that desegregation orders were considered temporary and should be lifted if a good-faith effort is made. Additionally, *Dowell* put forth that once unitary status is declared, districts are permanently freed from supervision and may revert to past student assignment practices. The following year, the Court further facilitated unitary status declarations through *Freeman v. Pitts* (1992) by ruling that districts can achieve unitary status on individual aspects of their desegregation plans separately as opposed to accomplishing the plan and its goals in their entirety. Lastly, in *Missouri v. Jenkins* (1995), the Court voted 5-4 to relocate the burden of proof of articulating racial disparities from the districts to their plaintiffs and moved the "finish line" from true racial balance to merely school racial compositions comparable to what they would have been if de jure segregation had not occurred. This served as tacit acceptance of school segregation stemming from residential segregation. Scholars regard these three cases as the end to hopes of full school desegregation, as "the treatment was too short and too incomplete and the forces supporting segregation in the housing markets, in schools, and elsewhere were far more resilient than the Court assumed" (Orfield & Lee, 2007, p. 9).

In 1998 and 1999, two U.S. District Courts and two U.S. Courts of Appeals published rulings that further hindered active pursuits of school desegregation. In these four cases, each Court struck down magnet program plans that considered race in students' magnet program applications (*Capacchione v. Charlotte-Mecklenburg Board of*

Education, 1999; *Eisenberg v. Montgomery County Public Schools*, 1998; *Tuttle v. Arlington County School Board*, 1999; *Wessmann v. Gittens*, 1998). Each plan used race as a weighting factor or tie-breaker to ensure that magnet enrollment patterns supported school desegregation efforts rather than increased racial isolation, but the Courts expressed concern over whether this represented a compelling interest.

In 2007, the U.S. Supreme Court addressed this question in the K-12 public education context in *Parents Involved in Community Schools v. Seattle School District No. 1* (PICS, 2007). In this case, the Court considered desegregation plans in school districts in Seattle, Washington and Louisville, Kentucky, which were no longer under court order but strove to further racial diversity of their own accord. These plans considered student race as a tie-breaker when demand exceeded availability for enrollment in the districts' magnet programs. The Justices had much to say, as evidenced by the five opinions published in this 5-4 ruling. The Court concluded that a school's racial diversity is a compelling state interest but that these race-conscious student enrollment policies were not narrowly tailored to accomplish this interest among schools not under a court order. While some magnet program scholars argue that this case represents "the culmination of several decades' worth of legal backpedaling on the implementation of *Brown*" (Siegel-Hawley & Frankenberg, 2012, p. 7), others assert that the ruling is "by no means … movement-ending" (Frankenberg & Le, 2008, p. 1015). Perhaps the most realistic perspective is that this case in fact "will have little practical effect on school districts because racial integration is off of the agenda of most school districts" (Holley-Walker, 2010, p. 879) already. Longitudinal research will be necessary to determine the true impact of these determinations on school racial segregation.

The End of ESAA

In the first year of his presidency, President Ronald Reagan championed the 1981 Omnibus Budget Reconciliation Act in an effort to curtail government spending (Christenson et al., 2003; Clinchy, 1993; Dentler, 1990; Frankenberg & Siegel-Hawley, 2010; Orfield & Lee, 2007). One facet of the act was a drastic cut in ESAA funding as the grant program was condensed into generic block grants (Blank et al., 1983; Christenson et al., 2003; Clinchy, 1993; Dentler, 1990; Frankenberg & Siegel-Hawley, 2010; Orfield & Lee, 2007). By doing so, the magnet program grant funding upon which many districts relied decreased from approximately $400 million in 1979 to $25 million in 1982 (Blank et al., 1983). In effect, this act "killed this revolutionary program [ESAA] and left beleaguered urban districts stranded" (Clinchy, 1993, p. 31). Considered together, this withdrawal of financial support for integration efforts combined with the dilution of desegregation mandates constituted a federal deprioritization of racial desegregation in public schools.

The Magnet Schools Assistance Program

In 1985, Congress enacted the Magnet Schools Assistance Program (MSAP), a federal grant program for districts seeking to implement or improve magnet programs at their schools (Christenson et al., 2003; Clinchy, 1993; Frankenberg & Le, 2008; Goldring, 2009). One possible contributing factor to MSAP's passage may have been the 1983 *Nation at Risk* report, which publicized the country's poor academic performance in comparison to foreign nations (Frankenberg & Le, 2008). Unlike ESAA, MSAP funds cannot be used to support any desegregation technique besides magnet programs (Magnet Schools Assistance Program, 1985). This grant program continues to this day.

MSAP Eligibility. In order to apply for MSAP funding, school districts must be under a court-ordered desegregation plan or develop a voluntary desegregation plan that is approved by the U.S. Secretary of Education (Magnet Schools Assistance Program, 1985; Steele & Eaton, 1996; Walton & Ford, 2014). These districts must also serve a substantial number of students of color (Magnet Schools Assistance Program, 1985).

MSAP Application Process. To apply for an MSAP grant, districts must provide current enrollment data, disaggregated by student group, for the district and all of its campuses, as well as disaggregated enrollment projections for each year of the grant period (Magnet Schools Assistance Program, 1985; Steele & Eaton, 1996). In their applications, districts must set a goal of reducing, eliminating, or preventing minority student isolation throughout the district, although no numerical targets are required (Magnet Schools Assistance Program, 1985). Minority student isolation is defined as students of color being in the majority of a school's enrollment when they are a minority of overall district enrollment (Magnet Schools Assistance Program, 1985; Walton & Ford, 2014). Later administrations amended the requirements to include additional goals, which will be described below in the "Changes to MSAP" section of this chapter.

MSAP Functionality. MSAP grants initially lasted for two years, until the grant was extended to a three-year term in 1995 (Christenson et al., 2003; Magnet Schools Assistance Program, 1985). Typically, the Department of Education awards grants to 30-50 districts in each grant cycle (Siegel-Hawley & Frankenberg, 2012). Each grant can include up to four million dollars per year per district, although grants are typically less (Magnet Schools Assistance Program, 1985; Wang & Herman, 2017). Districts may apply for grants multiple times and even renew previous grants (Magnet Schools

Assistance Program, 1985; Steele & Eaton, 1996).

MSAP funds can be used in a number of ways to develop, implement, and/or sustain magnet programs (Magnet Schools Assistance Program, 1985). In a study of 49 grantee districts, Siegel-Hawley and Frankenberg (2012) found that 100% of districts used MSAP funds to upgrade learning technology, 98% for professional development, 96% for curricula and/or a teaching method that would otherwise be unavailable, and 86% for theme-based offerings (such as STEM or arts enrichment).

MSAP Award History. Within its first decade, the MSAP granted a total of $955 million to 138 school districts (Steele & Eaton, 1996). As of 2017, the program has distributed a total of approximately three billion dollars (Wang et al., 2017). However, this is a lower funding rate than ESAA provided (Clinchy, 1993), and MSAP funding has remained stagnant and uncontrolled for inflation since its inception (Frankenberg et al., 2008; Rossell, 2005). Magnet and desegregation proponents alike have called for an increase in funding, particularly as magnet programs have become more prevalent.

Changes to MSAP. Just as the jurisprudence affecting magnet programs has changed over time, so have magnet funding policies. While the MSAP has been in place for 35 years, it has morphed significantly to align with the nation's evolving priorities. This change is visible through the expansion from one program goal to six. Initially, the program's only goal was to (1) eliminate, reduce, or prevent minority isolation in participating districts (Magnet Schools Assistance Program, 1985; U.S. Department of Education, 2017). Shortly thereafter, the Reagan administration introduced a second goal representative of national sentiment following A Nation at Risk: (2) increased academic knowledge and vocational readiness (Frankenberg & Le, 2008; U.S. Department of

Education, 2017). In 1994, the Department of Education under President Bill Clinton contributed two additional goals, further pulling focus from the program's initial purpose of addressing public school segregation: (3) district-wide systemic reform that enables all students to master learning standards and (4) innovation that promotes choice and diversity (Frankenberg & Le, 2008; U.S. Department of Education, 2017). Two final goals were incorporated in 2001 to correspond with President George W. Bush's No Child Left Behind Act: (5) operating at high performance levels as measured by standardized test scores and (6) equitable access to postsecondary readiness (Frankenberg & Le, 2008; U.S. Department of Education, 2017). Instead of desegregation representing the full focus of this essential grant program, this issue is now a mere sixth of the MSAP's priorities. According to Frankenberg and Le (2008), this change supports the notion that the "the federal government over the years has diluted the desegregative purpose of magnet schools, demanding with each reauthorization and modification of its grant program that these schools pursue numerous other educational reform objectives" (p. 1056) and that "these new goals are at best tangentially related, and sometimes entirely unrelated, to the original desegregation goal" (p. 1056).

Another significant change to MSAP policy came in 2004, when the Department of Education under President Bush mandated that districts applying for a grant must write their desegregation goal in race-neutral language (Betts et al., 2015; Frankenberg & Siegel-Hawley, 2010; Siegel-Hawley & Frankenberg, 2012, 2013). The Obama Department of Education strove to remedy this by guiding districts in strategies for reducing minority isolation while maintaining compliance with the 2007 *PICS* ruling (Siegel-Hawley & Frankenberg, 2012). This assistance resulted in grantee districts pursuing more overt diversity goals, reaching out to more diverse communities, and offering more transportation to magnet program students (Siegel-Hawley & Frankenberg, 2012).

A Wave of Unitary Status Declarations

Between 1954 and 1990, 1,057 districts implemented court-ordered desegregation plans, with the majority assigned between 1968 and 1978 (Johnson, 2012). Since the 1970s, the number of districts under court order has consistently declined (Lutz, 2011), and over 300 districts remained under order as of 2014 (Hannah-Jones, 2014).

The revocation of desegregation court orders - also known by the label “unitary status” - can be initiated by the school district, a judge, the local government, the federal government, or a third party (Holley-Walker, 2010; Johnson, 2011; Lutz, 2011; Orfield & Lee, 2007; Reardon et al., 2012). The end of a court-ordered plan often means “that the court [has] formally washed its hands of further involvement” (Clotfelter et al., 2006, p. 370), although some courts grant unitary status with stipulations. Even when communities wish to continue pursuing desegregation, courts can override school districts and dissolve their desegregation orders, as well as forbidding voluntary plans (Frankenberg & Lee, 2002; Orfield & Lee, 2007).

Unitary status declarations escalated dramatically after the 1990s trifecta of Supreme Court cases detailed above (Lutz, 2011). Simultaneously, the nation was beginning to suffer “integration fatigue” (Hannah-Jones, 2014; Orfield & Lee, 2007). This fatigue occurred as many White families and policymakers regarded existing levels of desegregation as sufficient, given the end of de jure segregation, and lost interest in continuing efforts to address de facto segregation. Within the context of this loss of concern among empowered stakeholders, many districts have been released from court order within the past 20 years with increasing frequency and lowered standards (Frankenberg & Siegel-Hawley, 2010; Holley-Walker, 2010). This has taken place despite the fact that, in the South, these districts still primarily consist of one-race schools (Holley-Walker, 2010). For example, a district in Alabama was declared unitary in 2000 despite developing zero desegregation policies, having a high school with a 90% Black student enrollment, and refusing to rename a school named for a founder of the Ku Klux Klan (Hannah-Jones, 2014). School districts that are currently under court order tend to have so little oversight that, according to a ProPublica investigation (Hannah-Jones, 2014), district officials often have never read their district's order, mistakenly believe they have been granted unitary status, or even are unaware that the order exists. Suffice to say, the era of court-ordered school desegregation appears to be drawing to a close, particularly in the South (Orfield et al., 2014).

Unitary Status Leads to Resegregation. Following a unitary status declaration, few school districts maintain explicit efforts to encourage racial diversity as "priorities and policies focus on expanding choice options, neighborhood schools, and quality improvement, replacing the previous emphases on judicial oversight and racial balancing" (Smrekar, 2009, p. 210). This tends to result in schools resegregating, such that school segregation can reinforce or even exceed residential segregation patterns (Arcia, 2006; Lutz, 2011; Orfield & Lee, 2007; Reardon et al., 2012; Smrekar, 2009). As school districts resegregate following unitary status, academic achievement tends to decrease in schools attended by increasing proportions of students of color, potentially widening achievement gaps (Frankenberg & Le, 2008; Lutz, 2011; Reardon et al., 2012).

Through large-scale, quantitative research, scholars have confirmed the concept of post-unitary status resegregation. In a national study, Lutz (2011) implemented a methodology that enabled a causal finding that the dismissal of court-ordered desegregation plans directly led to increased segregation, undoing an average of approximately 60% of the effects of these plans. The following year, Reardon et al. (2012) determined that 215 districts that had been released from their court orders between 1991 and 2009 became increasingly more segregated for up to 10 or 12 years following the declaration of unitary status. In addition, Reardon et al. (2012) found that resegregation was greatest in the South, in elementary grades, in large districts, in substantially Black districts, in areas with high residential segregation, and in districts that were less segregated before their release from the order (Reardon et al., 2012). It is worth noting that resegregation effects may lag by three years as districts phase out practices required by their former court orders, as indicated by Clotfelter et al.'s 2006 study of the 100 largest Southern and Border districts released from court order. Additionally, Orfield and Lee (2007) determined that over 33% of the desegregation Black students have experienced since *Brown I* had been reversed as of 2007. Researchers have found significant evidence of resegregation following unitary status declarations in studies of individual districts as well (Clotfelter et al., 2006; Frankenberg & Lee, 2002; Smrekar, 2009). Keeping in mind the immensely significant long-term benefits of court-ordered desegregation discussed above, unitary status declarations are likely to have monumental ripple effects.

Effects of Magnet Programs on Racial Desegregation in Public Schools

Given their potential impact on student enrollment across a district, magnet programs may shape the segregation or desegregation of a school district. This depends on adequate development, implementation, and commitment, however. According to Dentler (1990),

> magnet policies can be adopted alternatively as a "shell game" in order to create the appearance of desegregation; to introduce a stall or a stop in the course of litigation; to set up havens for selective subgroups of parents and students … Alternatively, magnet plans can revitalize equity in a district, contribute to the build-up of public confidence in the system, diversify opportunities for students, and redistribute both staff and students district-wide in desegregative ways. (pp. 75-76)

The research exploring the effect of magnet programs on schools' racial desegregation proves to be substantially mixed. In fact, approximately equal amounts of studies have reported positive effects, negative effects, and mixed or neutral effects. Each category is explored in kind below and should be kept in mind when considering the effects of magnet programs on student learning, as academics and equity issues such as desegregation are inherently intertwined.

Positive Effects of Magnet Programs on School Desegregation

In two large-scale studies and several rigorous, district-specific evaluations, researchers have identified racial desegregation in schools following the adoption of magnet programs. In particular, ESAA-funded magnets appeared to serve diverse student groups, such that the U.S. Department of Education labeled over two-thirds of these schools "fully desegregated" (Frankenberg & Le, 2008). It is worth noting that this effect did not necessarily address the districts as a whole, nor did the effect appear to continue under the MSAP (Frankenberg & Le, 2008). Frankenberg et al. (2008) surveyed 235 magnet teachers and administrators and determined that schools with stated desegregation goals were significantly less likely to be one-race schools than schools without desegregation goals, including schools that previously had goals but had since discontinued them. Additionally, Blank et al. (1996) found that, among 600 districts offering magnet programs, magnets appeared to "attract and enroll students from the nondominant ethnic group," (p. 167). Taken together, these large-scale findings demonstrated that when magnet programs set desegregation goals and received adequate funding, schools can become less racially segregated.

Rigorous evaluations of individual districts following the implementation of magnet programs have been more prevalent than large-scale studies. This may be because many districts evaluate their programs in order to justify continued expenditures to local or outside stakeholders. For example, Grooms and Williams (2015) analyzed 18 magnet programs in St. Louis Public Schools and found that the district's magnets were more diverse than its nonmagnet schools. However, as the district shifted to leveraging magnet programs for academic innovation and school choice rather than desegregation, these programs became increasingly racially segregated (Grooms & Williams, 2015). Yu and Taylor (1997) found that Cincinnati Public Schools and Metro Nashville Public Schools achieved near-perfect racial balance in their magnet programs, such that their student bodies were virtually identical to the district's racial composition as a whole. Three years into the implementation of its court-ordered magnet programs, 24 of the 26 magnet programs in Charlotte-Mecklenburg Schools met the court's desegregation target, as did all of the nonmagnet schools in the district, indicating a phenomenal district-wide impact on racial balance (*Magnet Assistance Program Enrollment and Participation Report*, 1995). In a study of five California cities, Rickles et al. (2004) found that elementary students enrolled in magnet programs attended schools that were less segregated than their neighborhoods, suggesting that these programs "have a potential to sever the nexus between residential and school segregation" (p. 4). Betts et al. (2006) and Koedel et al. (2009) established that the magnet programs in the San Diego Unified School District helped to desegregate the district as well as increase school-level diversity in terms of students' past academic achievement and parental education levels. In comparison, a different choice method in the district that did not have a desegregation goal led to an increase in segregation (Betts et al., 2006; Koedel et al., 2009). Connecticut offers a statewide interdistrict magnet system in which students may apply to a magnet program in any district, regardless of their home address. This system resulted in magnet students attending more desegregated schools than they had been assigned to in their zoned districts, which tended to be racially isolated (Bifulco et al., 2009). Lastly, a small, diverse district in Montclair, New Jersey achieved racial balance in all nine of its schools by adopting a magnet program at each and inviting students to select between them while ensuring even racial distribution (Clewell & Joy, 1990). This research suggests that desegregation is possible among certain magnets in certain contexts, although generalizability is inherently limited in district-specific research.

Negative Effects of Magnet Programs on School Desegregation

The research demonstrating negative effects of magnet programs on desegregation is led by large-scale evaluations of MSAP grantee districts. As of 1989, MSAP grantees were required to pursue two goals: (1) eliminate, reduce, or prevent minority isolation and (2) increase academic knowledge and vocational readiness (Magnet Schools Assistance Program, 1985; U.S. Department of Education, 2017). Reflecting the decrease in focus on school desegregation, only 37% of the 1989-1991 MSAP grantee districts specified a desegregation goal, primarily reducing minority isolation (Steele & Eaton, 1996). Of these few districts, only 48% met their goal by the end of the two-year grant cycle, and two years after the cycle ended, only 40% of districts had met the goal, indicating that certain districts had experienced an increase in segregation following the end of MSAP funding (Steele & Eaton, 1996). The grantee districts in fact had more minority-isolated schools as a whole at the conclusion of the grant, including 10% more schools with 80% or higher minority enrollment (Steele & Eaton, 1996). Frankenberg and Le (2008) performed a similar evaluation of the 2001-2004 MSAP cohort and found that 55 of the 92 grantee districts labeled “hyper-segregated” became more segregated over the course of the grant. It is worth noting that the 2001-2004 cohort was the first to navigate the MSAP’s full set of six goals, five of which do not address school diversity (Frankenberg & Le, 2008; U.S. Department of Education, 2017). Finally, Walton and Ford (2014) discovered that within the 2010-2013 MSAP cohort, the percentage of schools meeting or making progress towards their desegregation target fell over the course of the grant, especially among urban and Title I schools. All three large-scale evaluations indicate that MSAP funding has not mitigated school racial desegregation.

Additional large-scale analyses of magnet programs have revealed similar findings. In a survey of 236 magnet teachers and administrators, Frankenberg et al. (2008) determined that the magnet programs represented were becoming increasingly more segregated, seemingly through a loss of White students. Siegel-Hawley and Frankenberg (2012) analyzed National Center for Education Statistics data and found that half of the Black students enrolled in a magnet program attended a racially isolated (90% or greater) school. In comparison, 35% of Black students not enrolled in a magnet program attended a racially isolated school (Siegel-Hawley & Frankenberg, 2012). Together with the MSAP evaluation results, these studies indicate that many magnet programs have not fulfilled their potential for desegregation.

Lastly, when considering the impacts of magnet programs on desegregation, within-school effects must be considered, as "second-generation" segregation can emerge within seemingly diverse schools. "Second-generation" segregation refers to segregation at the classroom level within desegregated school buildings, an issue which I explore below in the section titled "Second-Generation Segregation."

Mixed and Neutral Effects of Magnet Programs on School Desegregation.

The plurality of research in magnet programs' effects on racial desegregation offers mixed or neutral results. These findings are prevalent in MSAP cohort studies, other large-scale analyses, and district-level evaluations.

Three MSAP studies have presented mixed findings. First, Steele and Eaton's (1996) evaluation of the 1989-1991 MSAP cohort, a report funded by the U.S. Department of Education, demonstrated a degree of nuance worth examining. Although only 48% of the districts that had a desegregation goal met that goal by the end of the two-year grant cycle, the achievement varied significantly depending on the goal type (Steele & Eaton, 1996). 44% of the grantee districts that focused on reducing minority isolation were successful, 33% of those striving to eliminate isolation were successful, and 73% of those aiming to prevent isolation were successful (Steele & Eaton, 1996). Seven years later, Christenson et al. (2003) conducted the Department of Education's next large-scale MSAP report, evaluating the 1998-2001 MSAP grantee cohort. Among these districts, 57% decreased their minority isolation over the course of the grant cycle, although for the vast majority of schools, the decrease was merely between one and five percent (Christenson et al., 2003). Districts with voluntary desegregation plans, elementary schools, whole-school magnets, schools with multiple minority groups, schools with high parental involvement, and schools with low student-to-teacher ratios proved to be the most successful in decreasing segregation (Christenson et al., 2003). Finally, Betts et al. (2015) analyzed 21 elementary magnets established through funding from a 2004-2007 or 2007-2010 MSAP grant and found mixed results as well. They specifically examined the effect of using grant funds to introduce a magnet program to a school that previously did not offer one and found that previously "disadvantaged" schools experienced an increase in racial diversity following the adoption of a magnet, while previously "advantaged" schools did not (Betts et al., 2015). According to these MSAP evaluations, magnets can produce mixed results in terms of school diversity, depending in part upon select characteristics of the school and district.

Studies of magnet programs beyond the scope of MSAP funding have indicated this variability as well. In their mixed methods national study of 45 magnets representing 15 urban districts, Blank et al. (1983) found that, while two-thirds of the magnet programs were racially balanced (reflecting the demographics of their districts as a whole), the teachers at these programs were disproportionately White. The researchers also developed an index of true racial integration, going beyond demographic statistics of desegregation to analyze information about each school's culture, community, and climate (Blank et al., 1983). They discovered that this integration index correlated with the percentage of students who were Black, principal quality, flexibility given by the district, adherence to magnet theme, and education quality (Blank et al., 1983). In 2014, Davis analyzed enrollment data from 1,437 schools. She found that schools with and without magnet programs within the sample had roughly equivalent racial diversity at the school level, despite some differences at the classroom level (Davis, 2014). These two studies' findings aligned with the previous results from MSAP analyses to some degree.

Lastly, mixed and neutral results have been found through district analyses as well. The Kansas City School District in Missouri implemented the "most ambitious and expensive desegregation program ever undertaken to entice students into city schools," (Morrison, 1996, p. 132) in which nearly every school in the district adopted a magnet program. This plan cost well over one billion dollars (Morrison, 1996). Despite this, the district fell short of its desegregation goal by 75% (Morrison, 1996). Although this initiative is credited with increasing student retention in the district to a slight degree, the district failed to meet its goal of increasing White enrollment following White flight of 26,000 students (Morrison, 1996). Chicago and Houston also implemented substantial magnet programs in response to White flight with mild results. In Chicago Public Schools, only eight of the 15 magnets included in a court order met their specified desegregation goals (Allensworth & Rosenkranz, 2000). In the Houston Independent School District, around half of the magnet programs developed during a 15-year initiative met their desegregation targets, due in part to the loss of 50,000 White students in that period (Stanley, 1989). Considering the research presented in this as well as the previous two subsections, one cannot reasonably conclude that magnet programs increase or decrease racial segregation within public schools, as the extant research offers such a high degree of heterogeneity.

Second-Generation Segregation

Although campus-level desegregation is a vital prerequisite for a diverse educational environment, it is not sufficient on its own. Classrooms must be desegregated as well for school desegregation to impact educational opportunities (Clotfelter, 2004; Darling-Hammond, 1997; Davis, 2014; Dentler, 1990; Frankenberg et al., 2008; Kahlenberg, 2001; Mickelson, 2001; Morris & Goldring, 1999; Oakes, 1986; Orfield & Lee, 2007; Rubin et al., 2006; Welner, 2006; West, 1994). "Second-generation" segregation refers to classroom-level segregation within desegregated school buildings, which may result directly from explicitly segregative policies or indirectly through academic tracking. Regardless of the source, second-generation segregation "serves to reproduce the larger social structure and the inequalities therein" (Davis, 2014, p. 5).

Second-generation segregation causes substantial harm. Many schools, primarily secondary schools (Clotfelter, 2004), offer academic tracks such as "regular," advanced, and honors. Different tracks provide vastly disparate opportunities to learn in terms of curriculum, resources, teacher quality, teacher experience and qualification, and cognitive engagement (Clotfelter, 2004; Darling-Hammond, 1997; Kahlenberg, 2001; Mickelson, 2001; Oakes, 1986; Rubin et al., 2006). Students of color and students living in poverty are substantially overrepresented in lower-track courses, in which they typically receive inferior opportunities, thereby widening gaps between student groups (Kahlenberg, 2001; Oakes, 1986). This serves to "unnecessarily buy the achievement of a few at the expense of many” (Oakes, 1986, p. 15).

The use of tracking to deliberately accomplish second-generation segregation of students by race was found unconstitutional in *Hobson v. Hansen* (1967). This practice had been used in a number of districts, particularly in the South, to fulfill court-ordered desegregation requirements while still sequestering White students (Clotfelter, 2004). A second federal case, *Hart v. Community School Board of Brooklyn* (1974), reinforced this ruling. Unfortunately, this practice was not abandoned in the 1970s. In 2001, the Seventh Circuit Court of Appeals determined that Rockford School District in Rockford, IL specifically implemented program-within-school magnets so that White students could attend entirely separate classes from their peers of color, who received a lesser curriculum (*People Who Care v. Rockford Board of Education School District No. 205*, 2001). In such situations, program-within-school magnets can enable “overall building desegregation by attracting enough white transfer students to balance the number of neighborhood minority students ... yet ... white transfer students rarely take classes with the minority nontransfer students” (West, 1994, pp. 2569-2570). This produces the veneer of desegregation while, within the school building, students remain segregated, and opportunity gaps remain rampant.

Three examples demonstrate the effects of second-generation segregation through tracking. Mickelson (2001) analyzed student-level data in Charlotte-Mecklenburg Schools, which was recognized as a highly desegregated district, and found that students of color were highly underrepresented in advanced academic tracks, allowing “Whites [to] retain privileged access to greater opportunities to learn” (p. 243). All else held equal, including prior academic achievement, placement in a lower track correlated with less rigor, inferior curriculum, less qualified teachers, lower end-of-course grades, lower GPA, and lower SAT scores among twelfth grade students (Mickelson, 2001). In Jacques (1993), a veteran high school counselor described tracking at magnet programs in her school district in Miami, Florida by saying, “Sometimes children are not even allowed to socialize with magnet students ... there's also favoritism, easy schedule changes for magnet students opposed to non-magnet students and very little mixing if any” (p. 12A). Lastly, a four-year case study of a large, diverse high school in California determined ninth grade math to be a major determinant of academic track, with tracks becoming more racially segregated as students progressed through high school (Rubin et al., 2006). This pivotal moment particularly harmed students with less-involved parents, as the school required students to self-select courses and teachers with insufficient assistance from counselors (Rubin et al., 2006). This manifested as a process “difficult to navigate without informed parental support, insider knowledge, and perseverance” that “[reinforced] existing patterns of racial segregation and [played] an important role in reproducing patterns of academic success and failure” (Rubin et al., 2006, p. 84). Parental lobbying can be a significant determinant in gaining access to rigorous tracks, and the fact that middle-class parents are most likely to feel comfortable exerting pressure on school officials means that such systems can contribute to second-generation segregation and subsequent opportunity gaps (Clotfelter, 2004). These cases represent how “the potential academic benefits of the 30 years of desegregation mandated by the *Swann* decision are compromised - even subverted - by the pervasive resegregation of secondary students into racially isolated tracked core academic classes” (Mickelson, 2001, p. 217).

Lastly, magnet researchers and practitioners should consider the nonacademic ramifications of second-generation segregation. Social segregation is a common feature within schools segregated at the classroom level (Clotfelter, 2004; Jacques, 1993; Noguera & Wing, 2006; Rubin et al., 2006). When students are deprived of opportunities to build relationships with peers in the classroom, they often lack the foundation to seek out cross-racial friendships or interactions in extracurricular activities (Clotfelter, 2004; Noguera & Wing, 2006; Rubin et al., 2006). These nonacademic contexts, in fact, can rapidly become more segregated than classrooms (Clotfelter, 2004; Noguera & Wing, 2006). In the California high school described above, which experienced high classroom-level segregation, only two of the school's 73 extracurricular and athletic groups were racially mixed (Rubin et al., 2006). Faculty advisors shared that they did not know how to recruit diverse students and/or felt reluctant to do so (Rubin et al., 2006). Students reported wanting to create spaces for themselves and those like them, to be with existing friends, to engage in activities that their friend group regarded positively, and/or to select activities based on their perceived benefit for college applications (Rubin et al., 2006). Considering this manner in which second-generation segregation can enable a feedback loop of racial isolation, district, campus, and magnet program leaders must keep in mind that "desegregation - by racially balancing schools - cannot be an objective, in and of itself; desegregation efforts have to be evaluated by the extent to which they result in equitable practices for all students" (Morris & Goldring, 1999, p. 64), academically and otherwise.

Magnet Programs Today

In comparison to charter schools, private schools, voucher programs, and homeschooling, magnet programs represent the largest sector of school choice in United

States education today, in terms of both number of programs and number of students served (Frankenberg et al., 2008; Goldring, 2009; Siegel-Hawley & Frankenberg, 2012, 2013). However, contemporary magnets tend to pursue different priorities than their predecessors of the 1960s through 1980s. This section details these priorities, as well as characteristics of today's magnet programs, their students, and their policies.

Mission of Contemporary Magnet Programs

Overall, twenty-first century magnet programs are widely considered to focus more on academic performance than desegregation and equity (Fleming, 2012; Frankenberg et al., 2008; Frankenberg & Le, 2008; Frankenberg & Siegel-Hawley, 2010; Goldring & Smrekar, 2000; Orfield & Lee, 2007; Rhea & Regan, 2007; Rossell, 2003; Siegel-Hawley & Frankenberg, 2012, 2013; Smrekar & Goldring, 2011; Wang et al., 2014; Yang et al., 2005). "Mission creep" (Frankenberg & Le, 2008) has occurred due to a combination of factors, including a more competitive school choice marketplace (Frankenberg et al., 2008; Frankenberg & Siegel-Hawley, 2010; Goldring, 2009; Wang et al., 2014), a call to serve as "incubators for innovation" (Fleming, 2012; Rossell, 2003; Smrekar & Goldring, 2011), an increased focus on high-stakes testing (Fleming, 2012; Frankenberg et al., 2008; Goldring, 2009; Rossell, 2003; Smrekar & Goldring, 2011), and a decreased national interest in racial desegregation in public schools (Betts et al., 2015; Fleming, 2012; Frankenberg et al., 2008; Frankenberg & Le, 2008; Hannah-Jones, 2014; Orfield & Lee, 2007; Rossell, 2003; Siegel-Hawley & Frankenberg, 2012; Smrekar & Goldring, 2011). Given these forces, districts must navigate "uncertain terrain for sustaining, much less increasing, racially diverse learning opportunities" (Frankenberg et al., 2008, p. 8).

Role of the MSAP in This Change. Since its inception, amendments to the MSAP have been a driving force in the shift in magnet program priorities. These amendments have broadened the scope of the grant program from one purpose - racial desegregation of schools - to six goals, five of which center on academic performance (U.S. Department of Education, 2017). Because of this, some scholars and education leaders fear that "in trying to expand the statute to provide for everything, it has come to stand for little if anything" (Frankenberg & Le, 2008, p. 1056).

The decreasing priority of the MSAP's original goal has been apparent as early as 1989, when grantee districts had two goals to manage. In evaluating the 1989-1991 MSAP cohort, the U.S. Department of Education determined that only 37% of districts awarded a grant even submitted a desegregation goal with their application (Steele & Eaton, 1996). Further, the Department of Education chose not to examine whether or not districts included desegregation goals in its 2003 report, although this report did show that 43% of grantee districts ended the three-year grant cycle with equal or greater racial isolation than at the start of the grant (Christenson et al., 2003; Frankenberg & Le, 2008; Siegel-Hawley & Frankenberg, 2012). A lack of desegregation leadership by the Department of Education and the sole federal grant program for magnet programs sets an example for the nation's public school systems, and this precedent has dramatically deprioritized racial desegregation.

Additionally, in 2004, the MSAP's application policies were altered so that any districts choosing to comply with the requirement of submitting a racial desegregation goal - a peculiar, oxymoronic concept in and of itself - must write the goal in race-neutral language (Betts et al., 2015; Wang & Herman, 2017). This policy was enacted three years prior to the *PICS* (2007) ruling (Wang & Herman, 2017). As districts confront increasing limitations (such as this policy) as well as heightened expectations from the federal government, their state education agency, and stakeholders, they must consider creative strategies to promote diversity while still acquiring federal funding, such as MSAP grants (Smrekar & Goldring, 2011).

Characteristics of Contemporary Magnet Students

As stated previously, the most recent reliable national data indicate that, as of 2017, close to 3.5 million students attend magnet programs in the United States (Magnet Schools of America, 2017). Wang and Herman (2017) analyzed enrollment statistics from 34 states compiled in the 2013-2014 Common Core of Data dataset and found that Hispanic, Black, and White students comprised roughly equal proportions of magnet program enrollment (33%, 27%, and 30% of enrollment, respectively). These findings were highly consistent with data from the 2008-2009 Common Core of Data (Siegel-Hawley & Frankenberg, 2013). For comparison, as of 2009, the nonmagnet student population in the United States was 22% Hispanic, 15% Black, and 57% White (Siegel-Hawley & Frankenberg, 2013). Additionally, in their 2017 research, Wang and Herman found that 59% of students in magnet programs in the 2013-2014 school year qualified for free or reduced-price lunch (FRPL). This compares to 45% of nonmagnet students qualifying for FRPL (Siegel-Hawley & Frankenberg, 2013). In their conclusion, Siegel-Hawley and Frankenberg (2013) argued that these statistics suggested that magnets are indeed serving the communities who have historically been denied enriched learning opportunities.

Characteristics of Contemporary Magnet Programs

At the program level, today's magnets offer a wide variety of formats, themes, and diversity guidelines (Siegel-Hawley & Frankenberg, 2013). While 23% of nonmagnet schools are in urban areas, 55% of magnets are located in urban communities (Wang & Herman, 2017). This number has been 10% higher among MSAP grantees (Walton & Ford, 2014). 73% of magnet programs are offered at schools receiving Title I federal funding to serve their high proportions of students living in poverty (Wang & Herman, 2017). The U.S. Department of Education's report describing the 2010 MSAP grant recipients indicated that five magnet program themes were particularly popular: science, technology, engineering, and math (38%); arts and humanities (22%); International Baccalaureate (17%); career and technical (7%); and foreign language and cultural studies (5%) (Walton & Ford, 2014). Additionally, Walton and Ford (2014) found that 32% of recipient schools had been labeled "improvement required" by the federal government or local education agency due to poor standardized test scores, and 23% of recipient districts still had court orders for mandatory racial desegregation in place. Most districts with magnet programs, according to Frankenberg and Siegel-Hawley (2010), offer them as a voluntary option within a school choice plan. So, while magnet programs display a wide variety of characteristics, most are urban, receive Title I funding, and are not under supervision for insufficient academic performance or racial balance.

Application Policies

School districts determine which students will attend magnet programs through a variety of methods. These include random lotteries, weighted lotteries (in which certain student groups are more likely to win, such as students who have a sibling at the school), open enrollment (or "first come, first served"), controlled choice (open enrollment within a limited subset of the district's magnets), interdistrict choice (in which students can transfer from their zoned district to a magnet at a different school district), and competitive selection processes (Siegel-Hawley & Frankenberg, 2012; Walton & Ford, 2014). In their analysis of the 2010 MSAP grant recipients for the Department of Education, Walton and Ford (2014) determined that 33% of magnet programs funded used a random lottery, 36% used a weighted lottery, 30% used open enrollment, and less than 1% used a competitive selection process. These percentages are somewhat skewed, as MSAP guidelines virtually prohibit the use of competitive selection processes among grant recipients (Frankenberg et al., 2008; Frankenberg & Siegel-Hawley, 2010). Competitive selection has historically been more prevalent among the vast amount of magnet programs that do not receive MSAP funding (Blank, 1989; Blank et al., 1996; Crain et al., 1992; Frankenberg et al., 2008; Siegel-Hawley & Frankenberg, 2013; Smrekar & Goldring, 2000).

However, school districts have a substantial degree of latitude in developing creative application processes as well. For example, in New York City Public Schools in the late 1980s and early 1990s, all eighth grade students in the district were required to complete a "high school application," selecting among a number of magnet programs as well as their zoned comprehensive high school (Crain et al., 1992). 82% of students in the district selected a magnet as their first choice (Crain et al., 1992). From there, the district implemented a shared enrollment protocol, in which magnet program administrators selected 50% of their new students and the other 50% were determined through a lottery (Crain et al., 1992). Further, both halves had a unique requirement - one sixth of the students in each half had to be among the top sixth of the district in terms of reading scores, and one sixth of each half had to represent the bottom sixth of reading scores (Crain et al., 1992). No matter the manner in which a magnet program selects its students, the policies must be narrowly tailored in order to withstand legal challenge (Welner, 2006).

Competitive Selection Processes. Competitive selection processes determine admission to magnet programs in a "meritocratic" manner. For example, a program might use a combination of students' test scores, GPAs, interviews, recommendations, auditions, attendance records, disciplinary records, specific course grades, and/or writing samples (Blank et al., 1996; Frankenberg et al., 2008; Siegel-Hawley & Frankenberg, 2013). The use of such criteria has increased as magnets have become more prevalent and less focused on desegregation (Frankenberg et al., 2008). The most recent rigorous data indicated that around one third of magnet programs nationwide use some form of selective admissions (Smrekar & Goldring, 2000). Interestingly, a survey of 236 magnet teachers and administrators found a correlation between competitive selection processes (especially test scores, auditions, and GPAs) and higher levels of racial segregation, as well as a correlation between the use of lotteries or open enrollment and lower levels of segregation (Frankenberg et al., 2008). Districts should thoroughly consider the implications of various student selection processes when developing magnet programs.

Demand and Waitlists

Historically, many magnets have experienced more demand for enrollment than availability, and excess demand appears to be increasing (Blank et al., 1996; Christenson et al., 2003; Goldring & Smrekar, 2002; Siegel-Hawley & Frankenberg, 2012). When programs experience a surplus of applicants, many districts develop a waitlist to enable additional students to enroll if seats become available. Blank et al. (1996) conducted the first large-scale research in this area and found that, in 1992, 53% of all magnet programs experienced more demand than supply, with over 123,000 students nationwide on a waiting list. Then, in their evaluation of the 1998-2001 MSAP cohort for the U.S. Department of Education, Christenson et al. (2003) determined that 63% of the cohort's magnet programs had surplus demand. A survey of MSAP-funded magnets roughly a decade later demonstrated an even higher percentage, with 73% of programs experiencing more demand than supply, such that 24% of survey respondents reported denying admission to between 1,500 and 7,500 students (Siegel-Hawley & Frankenberg, 2012). A final compelling finding is that magnet programs with increasing levels of racial diversity encounter more demand (Frankenberg et al., 2008). Considered in context, this growing pattern of excess demand for magnet enrollment proves particularly impressive given the continuously increasing numbers of magnet programs offered around the country (Blank et al., 1983, 1996; Blank & Archbald, 1992; Christenson et al., 2003; Frankenberg et al., 2008; Goldring, 2009; Siegel-Hawley & Frankenberg, 2012, 2013; Smrekar & Goldring, 2000; Steele & Eaton, 1996).

School Choice

Magnet programs are the most prevalent form of school choice in the United States, compared to other options including charter schools, private schools, homeschooling, and voucher programs (Frankenberg et al., 2008; Goldring, 2009; Siegel-Hawley & Frankenberg, 2012, 2013). To contribute additional context to this analysis of contemporary magnet programs, it is worthwhile to examine the broader concept of school choice.

The notion of school choice has legal roots that date back nearly a century in the United States, but this idea has become particularly prevalent in recent years, now pervading national conversations about education policy. Researchers have analyzed the impact of a number of factors on this issue, including considerations in families' decision-making and access to information. The data reveal that race and equity underpin a number of issues related to school choice. Above all, the research indicates that "the process of choosing schools is complex and contextually bound" (Smrekar, 2009, p. 213). As magnet programs represent the largest sector of school choice in the United States (Frankenberg et al., 2008; Goldring, 2009; Siegel-Hawley & Frankenberg, 2012, 2013), the processes, mechanisms, and equity concerns presented here are highly relevant to the broader discussion of how magnets impact opportunity gaps.

The Roots of School Choice

Pierce v. Society of Sisters (1925) first codified school choice in the United States, when the U.S. Supreme Court affirmed a liberty right among parents and guardians to select a school for their children. This precedent has enabled families to opt into private schools for generations, but only within the past 50 years has an expansive market for school choice among public schools developed. Magnet programs, voucher systems, and charter schools offer families alternatives to the traditional, zoned public schooling. Choice advocates argue that zoned school assignment "began as an administrative expedient without roots in pedagogical theory" (Doyle & Levine, 1984, p. 265) and that the paradigm shift towards school choice is an opportunity for U.S. public education to better serve students, a stance grounded in market theory.

Market Theory

Market theory applies economic concepts - including supply, demand, competition, and innovation - to the context of public schools (Chubb & Moe, 1990). According to this theory, families act as consumers who weigh strengths and weaknesses of school options for their children and select accordingly, bringing their children's per-pupil funding allotments with them in many states (Berends, 2014; Chubb & Moe, 1990). This autonomy is theorized to drive schools to compete for students and their funding, compelling them to innovate and enhance their educational offerings (Chubb & Moe, 1990). In doing so, market theory suggests that the quality of public education as a whole will improve due to this competition (Chubb & Moe, 1990).

Market theory principles are apparent within many districts' magnet programs. Keeping customers satisfied can supersede other intentions, including mitigating school segregation. In many cases, "school districts with choice plans have given priority to the parental-autonomy (free market) justification, making diversity a secondary goal at best" (Welner, 2006, p. 365). Additionally, students and families can also wield power by lobbying for the addition, expansion, or removal of a magnet program, including advocating for particular magnet themes (Blank et al., 1996). The public school choice movement has empowered students and families in an entirely new way.

Participating in Choice

Not all families and communities exercise school choice at equal rates. In fact, quantitative researchers have revealed two consistent patterns: families that engage in school choice tend to have higher incomes (Engberg et al., 2011; Goldring & Hausman, 1999; Martinez et al., 1994, 1996) and greater parental education attainment (Betts et al., 2006; Engberg et al., 2011; Martinez et al., 1994, 1996). Additionally, choice appears less prevalent among families whose parents are underemployed (Martinez et al., 1994) and families whose children are English learners (Betts et al., 2006). Taken together, these descriptors suggest that families who are more “advantaged” (Engberg et al., 2011) tend to disproportionately capitalize on school choice opportunities.

Researchers have identified contradictory results regarding the correlation between family race and school choice participation, suggesting that this varies across contexts. In St. Louis, Goldring and Hausman (1999) determined that White families were most likely to choose a school, while Betts et al. (2006) found that Black families participated in school choice in the San Diego Unified School District more than other racial groups. This supports Smrekar’s (2009) conclusion that school choice is contextually bound.

What Factors Families Consider in School Choice

Before detailing the existing research regarding factors families weigh when choosing schools, it is worth noting two potential limitations: response bias and issues related to rational choice theory. First, surveys serve as the primary methodology in this field. Schneider and Buckley (2002) argue that “relying simply on survey data to find out how parents will exercise their expanding rights to choose can lead to an overly optimistic view of what will motivate their actual choices” (p. 142). Parents may misrepresent their choice rationales on surveys and in focus groups to align with “socially desirable responses” (Teske et al., 2007, p. 22), a phenomenon known as response bias. To address this theory, Schneider and Buckley (2002) tracked parents’ Internet research when exploring school choice options and found that parents primarily investigated the demographics of schools' students, not academic information. Second, choice research has historically been grounded in rational choice theory, a branch of market theory which postulates that consumers rationally weigh the costs and benefits of their options and select the most objectively advantageous avenue (Smrekar & Goldring, 2000). However, this theory has become defunct within the field of economics, in part due to the work of psychologist Dr. Daniel Kahneman, who won the Nobel Prize in Economics for determining that human beings are often irrational in decision-making (Smith, 2002). Both of these findings - that individuals may be dishonest, and individuals may be irrational - should be kept in mind while reviewing the following research.

Through surveys, researchers have identified four primary factors that families consider in school choice, particularly with regard to magnet programs (Bauch & Goldring, 1995; Blank, 1989; Hausman & Goldring, 2000; Martinez et al., 1994; Smrekar, 2009; Smrekar & Goldring, 2000; Teske et al., 2007; Yu & Taylor, 1997), as well as several secondary considerations. In order of prevalence, the four key factors are academics, proximity and transportation, values, and discipline and safety. First, academic features including academic reputation, teacher quality, curriculum, and principal reputation are the most common rationale parents list in describing their choice process (Bauch & Goldring, 1995; Hausman & Goldring, 2000; Martinez et al., 1994; Smrekar, 2009; Smrekar & Goldring, 2000; Teske et al., 2007; Yu & Taylor, 1997). This factor is heavily valued among families of color (Teske et al., 2007), White families (Yu & Taylor, 1997), and higher-income families (Martinez et al., 1994; Smrekar & Goldring, 2000; Yu & Taylor, 1997) alike. Second, families consider the proximity and convenience of school choice programs, including transportation offerings (Bauch &

Goldring, 1995; Blank, 1989; Hausman & Goldring, 2000; Martinez et al., 1994; Smrekar, 2009; Smrekar & Goldring, 2000; Teske et al., 2007; Yu & Taylor, 1997). In particular, lower-income families (Martinez et al., 1994; Teske et al., 2007; Yu & Taylor, 1997) and families of color (Bauch & Goldring, 1995; Smrekar & Goldring, 2000; Yu & Taylor, 1997) weigh accessibility to schools. An alignment of family and school values or morals represents a third factor frequently identified by researchers (Bauch & Goldring, 1995; Hausman & Goldring, 2000; Smrekar, 2009; Smrekar & Goldring, 2000; Teske et al., 2007; Yu & Taylor, 1997). This holds true for both families of color (Yu & Taylor, 1997) and lower-income White families (Teske et al., 2007). Finally, student discipline and school safety have been considered together as the fourth major factor (Bauch & Goldring, 1995; Hausman & Goldring, 2000; Smrekar, 2009; Yu & Taylor, 1997). Lower-income families in particular report this influencing their decision-making (Bauch & Goldring, 1995; Yu & Taylor, 1997). Far and away, these four factors represent the most consistently self-reported variables families use in selecting among school choice options.

Additionally, through surveys, researchers have identified four additional variables that appear less frequently but may still be worth considering. Frustration with a child’s zoned school appears to have motivated families to opt into school choice in San Antonio and St. Louis (Martinez et al., 1994; Smrekar & Goldring, 2000). Additionally, families of color and families of lower socioeconomic status in Cincinnati, St. Louis, and Nashville valued the degree to which a school could provide specialized academic help for their child (Yu & Taylor, 1997). The curriculum or magnet theme offered by a school influenced lower-income families in Milwaukee, Washington D.C., and Denver (Teske et al., 2007). Last, families of color in Cincinnati, St. Louis, and Nashville reported seeking out schools with more opportunities for parental involvement (Yu & Taylor, 1997). Although these four factors are not widespread in the literature, they may still impact the decisions made by families participating in school choice.

Finally, a number of studies, including several with rigorous quantitative methodologies, have indicated that the demographics of a school's student body factor into some families' consideration of the school (Arcia, 2006; Betts et al., 2006; Henig, 1990; Holme, 2002; Saporito, 2003; Smrekar, 2009; Yu & Taylor, 1997). In fact, "perceptions of school quality are often tainted by race and class" (Frankenberg & Le, 2008, p. 1039). Most of this research has demonstrated an in-group/out-group effect, in which families leave schools where their children are in the racial or socioeconomic minority and opt into schools in which their children will be in the majority (Arcia, 2006; Henig, 1990; Holme, 2002; Saporito, 2003; Smrekar, 2009). Of the five studies just cited, all five analyzed schools' application and enrollment data, circumventing the possibility of response bias in a field typically characterized by surveys. Specifically, researchers have determined that White families were unlikely to apply to magnet programs in minority neighborhoods (Arcia, 2006), that White families preferred magnet programs that enrolled fewer students of color (Henig, 1990), and that families of color preferred programs in minority neighborhoods (Henig, 1990). Additionally, researchers have found that White families and non-poor families applied to magnet programs to leave their zoned school as their attendance zone became less White and/or poorer (Saporito, 2003). In fact, Smrekar (2009) identified that that once schools "tipped" to primarily enrolling students of color, White families left "with increasing and unbreakable momentum" (p.

222). In addition to enrollment analyses, Holme (2002) also interviewed 42 high-status parents. The families shared that they chose "good" schools because the parents regarded their children as gifted and worried that students of lower academic ability and with poor behavior would detract from their children's learning at "bad" schools (Holme, 2002). Holme (2002) examined the demographic data for the "good" and "bad" schools (as labeled by the interviewees) and found a substantial correlation with race, concluding that "by equating children of color with low academic achievement, these parents were able to express their concerns about diversity not in terms of racial or class prejudice, but in terms of concerns about the academic and social needs of their own children" (p. 195). Finally, research in the San Diego Unified School District showed that families of all races and socioeconomic backgrounds applied to schools with higher-achieving students, fewer English learners, and more highly educated parents (Betts et al., 2006). Considering these six studies together, it appears that school choice decisions are not as simple as seeking out a nearby school with strong academics.

What Information Sources Families Use in School Choice

Although school districts typically provide extensive information about school choice options through numerous modalities (described in the following section "School Choice Outreach to Families"), families rarely select a school for their child based on this information alone. In fact, "parents rely on multiple sources of information, trust word-of-mouth networks more than documentation, and, when push comes to shove, rely on other parents more than on teachers or administrators" (Teske et al., 2007, p. 39). Social networks appear to be the primary source of information in school choice decisions, including selecting a magnet program (Frankenberg & Le, 2008; Holme, 2002; Smrekar, 2009; Smrekar & Goldring, 2000; Teske et al., 2007). Typically, these networks offer insider information about schools' cultures, environments, ideologies, and reputations (Frankenberg & Le, 2008; Holme, 2002; Teske et al., 2007). Families of all backgrounds take advantage of their networks' knowledge, but middle- and upper-class families may have access to more valuable networks, including magnet teachers and connections to school boards (Holme, 2002; Smrekar & Goldring, 2000; Teske et al., 2007). By gaining information from social contacts, families of all backgrounds may garner "soft data" about prospective schools that is often valued above district-provided information.

Although social networks represent the predominant source of school choice information, families do utilize information provided by school districts. Teske et al. (2007) surveyed 800 low-income families who engaged in school choice in Milwaukee, Washington D.C., and Denver. 85% of parents visited a prospective school, 77% spoke with school officials, 74% had their children visit the school, and 73% read printed information provided by the district (Teske et al., 2007). As Teske et al.'s (2007) work centered on the decision-making process of lower-income families, it did not generate data to compare whether higher-income families utilize these options in addition to their social networks.

School Choice Outreach to Families

To ensure that the entire community is aware of school choice options and sufficiently informed to exercise them, districts must engage in deliberate outreach efforts. This is particularly relevant when districts aim to increase learning opportunities for historically underserved student groups. Without active efforts to engage prospective families who could benefit from participating in school choice, families with insider knowledge and well-connected social networks stand to benefit more from choice programs, further advantaging these families and reinforcing existing opportunity gaps (Holme, 2002; Smrekar & Goldring, 2000; Teske et al., 2007).

Recruitment methods vary between districts, but the most common are brochures provided to students and families, distributing applications directly to students, advertisements via local media (radio, newspaper, and television), school tours, information and applications mailed to students' homes, presentations at school fairs, flyers attached to report cards, brochures posted in local businesses and neighborhood associations, information booths at community events, and parent information centers to guide families through the choice process (Blank et al., 1996; Christenson et al., 2003; Frankenberg et al., 2008; Martinez et al., 1994; Yu & Taylor, 1997). Typically, districts offering magnet programs use multiple strategies to ensure thorough dissemination of information (Blank et al., 1996; Christenson et al., 2003, p. 1996; Yu & Taylor, 1997).

However, as described above, disparities exist across families as to the availability and depth of knowledge about choice options, with lower-income families potentially less able to rely on influential social networks for valuable knowledge (Holme, 2002; Smrekar & Goldring, 2000; Teske et al., 2007). To account for this gap, the onus lies on districts to provide information and assistance to lower-income families to ensure that school choice does not become a commodity limited to "advantaged" families. Yu and Taylor (1997) warned that "without thoughtful and carefully implemented outreach to poor families, this stated goal of choice programs [educational opportunity] may remain a paper promise" (p. 26). In his concurring ruling in *PICS* (2007), Justice Kennedy specified that outreach targeting families of certain races is a permissible method for magnet programs to pursue desegregation. MSAP grantees during George W. Bush's presidency rarely employed this option, but Obama-era grant recipients have been found to frequently and strategically reach out to diverse families (Siegel-Hawley & Frankenberg, 2012). Unsurprisingly, research has determined that districts and schools with explicit outreach efforts to diverse families have accomplished more desegregation than those without (Frankenberg et al., 2008; Siegel-Hawley & Frankenberg, 2013). In fact, these districts and schools experience greater demand from families of all racial groups (Frankenberg et al., 2008). "Lack of awareness regarding school choice programs appears to be a formidable obstacle to participation among some low-income minority families," (Martinez et al., 1994, p. 680) so targeted outreach is vital to ensuring equitable access to school choice, particularly magnet programs.

Opportunity Gaps

Racial opportunity gaps exist and persist in American public education (Alemán, 2006; Aud et al., 2011; Berends, 2014; Darling-Hammond, 1997; Gilliam, 2016; Gilliam et al., 2016; Gregory et al., 2006; Haycock, 2001; Kahlenberg, 2001; Ladson-Billings, 2010; Lee, 2002; Manning & Kovach, 2002; Mattison & Aber, 2007; Mickelson, 2001; Milner, 2010; Noguera & Wing, 2006; Oakes, 1986; Orfield et al., 2008; Orfield & Lee, 2007; Petras et al., 2011; Rothstein, 2004; Rubin et al., 2006; Scherff & Piazza, 2009; Skiba et al., 2002, 2011; U.S. Department of Education, Office of Civil Rights, 2014; Wang, 1998; Williams, 2002a, 2002b; Wing, 2006; Wu et al., 1982). In his majority ruling representing a unanimous U.S. Supreme Court in *Brown v. Board of Education* (1954), Justice Earl Warren wrote, "it is doubtful that any child may reasonably be expected to succeed in life if he is denied the opportunity of an education" (p. 493).

Given the historical purpose of magnet programs as tools for desegregation and their current reality as vehicles for academic enrichment, the interplay of race and academic opportunity contextualized in the fraught history of the racial opportunity gap provides valuable background for evaluating the degree to which magnet programs actualize their potential. Little research explicitly addresses opportunity or achievement gaps within magnet programs, so, in this section of the literature review, I offer a broader background on opportunity gaps to provide valuable context in the absence of more narrow, focused research.

A Shift in Language

Historically, the term "racial achievement gap" has been the conventional descriptor of the racial disparities in educational experiences and outcomes in U.S. public education. In his seminal 2010 book *Start Where You Are, But Don't Stay There*, Milner proposed rebranding this phenomenon as the racial "opportunity gap" (Ladson-Billings, 2010; Milner, 2010). Milner (2010) argued that achievement gap language can ascribe blame to students for inequities rather than the "systems, processes, and institutions [that] are overtly and covertly designed to maintain the status quo and sustain depressingly complicated disparities in education" (p. 8). More than semantics, he advocated for this shift in language to promote a paradigm shift to concentrate on inputs rather than outputs, striving to "[understand] reasons that undergird disparities" (Milner, 2010, p. 8).

Throughout this thesis, I use the term "achievement gap" only in reviewing research in which the data and results are limited to outcomes, particularly research indicating historical measures of achievement data. For discussion of inputs or personal commentary, I will employ the term "opportunity gap" to support this transition from "the exhausted 'achievement gap' lament to a more robust and nuanced discussion of why school failure persists for some groups of students" (Ladson-Billings, 2010, p. x).

Historical Measures of the Racial Achievement Gap

The Gap Narrows. The 1970s and 1980s were marked by decreasing racial segregation between Black and White students, primarily driven by court-ordered desegregation plans, a number of which mandated the implementation of magnet programs (Clotfelter, 2004; Reardon & Owens, 2014). Simultaneously, the Black/White achievement gap - as measured by performance on standardized tests and graduation rates - narrowed dramatically (Aud et al., 2011; Haycock, 2001; Lee, 2002; Manning & Kovach, 2002; Orfield & Lee, 2007; Williams, 2002a, 2002b). Racial gaps in both National Assessment of Educational Progress (NAEP) and Scholastic Aptitude Test (SAT) scores decreased during this time as well (Lee, 2002). Increased performance by Black students drove this reduction in achievement disparities, as White students' test scores remained stagnant (Lee, 2002). In fact, Black students demonstrated four times the annual growth of their White peers during this era (Lee, 2002). A correlation between desegregation and narrowing of the achievement gap is evident (Manning & Kovach, 2002). Lastly, it is worth noting that the achievement gap between White and Hispanic students changed inconsistently during this time, which researchers such as Lee (2002) have regarded as unsurprising, given the lack of a concerted focus on Hispanic/White school desegregation.

The Gap Widens. Beginning in the late 1980s, the Black/White achievement gap began to increase (Haycock, 2001; Lee, 2002; Manning & Kovach, 2002; Orfield & Lee, 2007; Williams, 2002a, 2002b). Researchers have established a correlation between this increase and a simultaneous increase in racial segregation in schools (Orfield & Lee, 2007). Lee (2002) identified 1988 as the inflection point for the racial disparity in NAEP and SAT scores. Additionally, Aud et al. (2011) identified an increase in the racial gap in dropout rates and correlated this change with the wave of unitary status declarations in the late 1980s in their large-scale analysis for the National Center for Education Statistics (NCES). The Hispanic/White achievement gap again did not follow the same pattern, remaining approximately stagnant from the early 1980s to the twenty-first century (Lee, 2002). Unfortunately, "although everybody wanted to take credit for narrowing the gap, nobody wanted to take responsibility for widening it. So, for a while, there was mostly silence" (Haycock, 2001, p. 6).

The Gap Persists. The gap in educational experience between White students and students of color was not left behind in the twentieth century; rather, it persists (Aud et al., 2011; Berends, 2014; Gregory et al., 2006; Haycock, 2001; Kahlenberg, 2001; Ladson-Billings, 2010; Lee, 2002; Manning & Kovach, 2002; Mattison & Aber, 2007; Mickelson, 2001; Milner, 2010; Noguera & Wing, 2006; Orfield et al., 2008; Orfield & Lee, 2007; Rothstein, 2004; Rubin et al., 2006; Scherff & Piazza, 2009; Williams, 2002a, 2002b; Wing, 2006). In their 2011 report for the NCES, Aud et al. identified a great deal of consistency in this finding. Specifically, gaps in Black/White high school graduation, Hispanic/White high school graduation, Black/White immediate college enrollment, Hispanic/White immediate college enrollment, Black/White college course rigor, Hispanic/White college course rigor, Black/White bachelor's degree attainment, and Hispanic/White bachelor's degree attainment have increased since the early 1990s (Aud et al., 2011). Just as the United States has not become a "post-racial" society, U.S. public education has not become a “post-opportunity gap” system. Much work remains to be done to reverse the expansion of these disparities, as current practices appear to be maintaining an inequitable status. This fact must be kept in mind by school district leaders as well as magnet program educators, particularly given the decreased focus on racial objectives within magnet programming over time.

Sources of the Racial Opportunity Gap

Misconceptions. Misunderstandings regarding the sources of the racial opportunity gap represent some of the greatest barriers to its mitigation. Historically, this gap has been blamed on family, cultural, and student factors, including poverty, motivation, academic effort, familial educational attainment, alcohol and drug use, and crime (Haycock, 2001; Lee, 2002; Noguera & Wing, 2006). Educators, unlike students, have been found to ascribe disparate student outcomes to these theories (Haycock, 2001). However, quantitative researchers have refuted such attributions, as the narrowing and widening of the achievement gap does not correlate with a similar pattern among any of these student-level variables (Lee, 2002). Noguera and Wing (2006) postulated that some educators may offset accountability for student learning as a defense mechanism. In fact, “by attributing the cause of minority student underachievement to a lack of student effort or deficient family background, we can comfortably dismiss the problem as sad and disturbing, and reject the possibility that something more pernicious might be at work” (Noguera & Wing, 2006, p. 6).

Evidenced Sources. What schools and districts do shapes racial opportunity gaps (Alemán, 2006; Darling-Hammond, 1997; Gilliam, 2016; Gilliam et al., 2016; Gregory et al., 2006; Guiton & Oakes, 1995; Haycock, 2001; Lee, 2002; Manning & Kovach, 2002;

Mattison & Aber, 2007; Mickelson, 2001; Milner, 2010; Oakes, 1986; Orfield et al., 2008; Petras et al., 2011; Skiba et al., 2002, 2011; U.S. Department of Education, Office of Civil Rights, 2014; Williams, 2002; Wing, 2006; Wu et al., 1982). One key finding supporting this conclusion is that, all external factors held equal, the longer a student of color is in school, the greater the gap between them and their White peers (Manning & Kovach, 2002). This finding supports the theory that the American education system "was never designed to teach *all* children effectively, to teach learners in all their varieties, to attend to each child's particular mix of aptitudes and barriers" (Darling-Hammond, 1997, p. xi, emphasis in original). Magnet programming offers a potential avenue to engage in this vital work, but without identifying and addressing the inequities structured into American public education, "status quo" practices will only reinforce the racial opportunity gap.

The greatest contributor to the racial opportunity gap is arguably the inferior allocation of material, financial, and human resources to schools that primarily serve students of color (Alemán, 2006; Darling-Hammond, 1997; Haycock, 2001; Lee, 2002; Manning & Kovach, 2002; Mickelson, 2001; Orfield et al., 2008; Williams, 2002; Wing, 2006). Traditional school finance schemes, in which local property taxes substantially determine school revenues, inherently advantage districts populated by middle-class, White residents (Alemán, 2006; Darling-Hammond, 1997). Although some states have reformed education funding to minimize these disparities, others have not (Alemán, 2006). Due in large part to these financial inequities, schools with high proportions of students of color often have fewer and older material resources, fewer certified teachers, less educated teachers, teachers with less teaching experience, higher student-teacher ratios, higher teacher turnover, less rigorous course offerings, and less rigorous instructional practices (Haycock, 2001; Lee, 2002; Manning & Kovach, 2002; Mickelson, 2001; Orfield et al., 2008; Williams, 2002; Wing, 2006). The U.S. school system "[takes] the students who have less to begin with and then systematically [gives] them less in school. In fact, we give these students less of everything that we believe makes a difference" (Haycock, 2001, p. 8). On the contrary, even controlling for socioeconomic status, White students typically enjoy more and newer resources, more qualified teachers, more advanced curricula, greater parent advocacy and involvement, and greater leniency with behavior and attendance infractions (Mickelson, 2001; Wing, 2006). Student interviews have shown that students interpret these disparities as an indication that schools and society value them unequally based upon race (Darling-Hammond, 1997). One of the factors undergirding hope in magnet programs is, in fact, the supplemental funding often provided to magnets, perhaps channeling greater resources to students of color than they would otherwise receive. Unsurprisingly, differences in educational inputs based on student demographics impact outcomes, including student achievement (Lee, 2002; Manning & Kovach, 2002; Orfield et al., 2008; Williams, 2002).

In addition to differing quantities and qualities of resources, disparities in expectations and discipline also play a role in opportunity gaps. Educators, particularly White educators, have been found to hold lower academic expectations for students of color compared to their White peers, which "contributes to an unending cycle: educators do not teach with rigor and high expectations; students do not learn; students' test scores suffer; and then all involved wonder why" (Milner, 2010, p. 36). Unequal expectations likely play a role in the racial discipline gap as well (Gilliam, 2016; Gilliam et al., 2016;

Gregory et al., 2006; Mattison & Aber, 2007; Petras et al., 2011; Skiba et al., 2002, 2011; U.S. Department of Education, Office of Civil Rights, 2014; Wu et al., 1982). Consistently, African-American and Hispanic students, especially boys, are overrepresented in referrals, suspensions, and expulsions compared to White students accused of the same infractions, particularly with subjective infractions such as defiance (Gilliam, 2016; Gilliam et al., 2016; Gregory et al., 2006; Petras et al., 2011; Skiba et al., 2002, 2011; U.S. Department of Education, Office of Civil Rights, 2014; Wu et al., 1982). Recent research suggests that teachers' implicit biases against students of color may be a significant factor in discipline disparities (Gilliam, 2016; Gilliam et al., 2016). Regularly anticipating poor academic and behavioral performance and punishing misbehavior more harshly results in students of color being "pushed farther and farther away from the school, and in the process they become further removed from the opportunities that [an] ... education is supposed to provide" (Gregory et al., 2006, p. 133). So, one cannot disregard issues of expectations and student discipline in seeking to understand and address the racial opportunity gap in American public education.

Finally, segregation - both school-level and second-generation segregation - in and of itself creates an opportunity gap (Guiton & Oakes, 1995; Mickelson, 2001; Orfield et al., 2008; Swanson, 2004). To be clear, it is a "nonsensical notion that desegregation enhances opportunities to learn via proximity to a diversity of derma" (Mickelson, 2001, p. 241). Rather, the primary benefit of having White peers is indirect, as resources, teacher quality, and family advocacy often follow these students (Mickelson, 2001). For this reason, magnet advocates may lament how desegregation has been greatly deprioritized within magnet programs, weakening one of the great assets they have historically offered: diverse classrooms. Tracking and second-generation segregation, as described above, widen opportunity gaps by denying students of color the advantages of a diverse learning environment, even when a school building is desegregated (Guiton & Oakes, 1995; Oakes, 1986). To be sure, "issues of power and privilege contribute to the continuation of poor schooling for many children" (Darling-Hammond, 1997, p. xvi) by structuring opportunity gaps into the ways schools and districts operate, including siloing the social and political capital of "advantaged" students.

Outcomes of the Racial Opportunity Gap

Given the pervasive elements of the racial opportunity gap in educational inputs, it should not be surprising that the outcomes resulting from this gap are extensive as well. Disparities in test scores, high school graduation, and education attainment reflect this most visibly (Aud et al., 2011; Harris, 2015; Mickelson, 2001; Orfield & Lee, 2007; Swanson, 2004). Students of color systematically receive fewer opportunities than their White counterparts, yet "we wonder why populations of students who have been poorly served in education cannot necessarily transcend our poorly run educational system" (Milner, 2010, p. 32).

Racial opportunity gap outcomes are particularly pronounced in secondary and postsecondary education. As described previously, Aud et al. (2011) conducted a rigorous analysis of national education data for the U.S. Department of Education NCES in 2011, detailing outcomes of the racial opportunity gap. Aud et al. (2011) found that White students outscored Black students by 26-27 points and Hispanic students by 22-25 points on the NAEP reading assessment. In math, the gaps were even greater, as White students scored 26-32 points higher than Black peers and 21-26 points higher than Hispanic peers

(Aud et al., 2011). White high school graduates were 8% more likely than Black peers and 9% more likely than Hispanic peers to matriculate to college immediately after graduating (Aud et al., 2011). Once in college, Black students were 14% more likely to enroll in remedial courses during their freshman year than White freshmen, and Hispanic students were 12% more likely to do so (Aud et al., 2011). Students of color experienced substantially greater college dropout rates, as White individuals were 19% more likely to attain a bachelor's degree than Black students and 25% more likely than Hispanics (Aud et al., 2011). In an additional study striving to provide further context to high school graduation data, Swanson (2004) calculated a "cumulative promotion index," aggregating the odds of promotion to tenth grade, promotion to eleventh grade, promotion to twelfth grade, and graduation. This alternate methodology revealed a 25% gap between Black and White students as well as a 22% gap between Hispanic and White students (Swanson, 2004). Both studies demonstrated that, after over a decade of schooling, students of different racial groups are very differently positioned to succeed in American society.

Although the bulk of the research exploring the outcomes of the racial opportunity gap has centered on high school students, significant consequences of such gaps are also apparent at the elementary level (Mickelson, 2001). Mickelson (2001) isolated the effect of attending a racially isolated Black elementary school and found it to correlate with lower sixth grade achievement scores, high school track, twelfth grade achievement scores, high school GPA, and SAT score. Opportunity gaps produce pernicious effects with the power to maintain the status quo, and this cycle does not wait until high school to begin. This emphasizes the needs for integrated, academically enriched learning environments at all levels, as magnet programs ideally offer.

Strategies to Mitigate the Racial Opportunity Gap

No silver bullet can eradicate the racial opportunity gap, including magnet programs. According to Williams (2002), "generic restructuring frameworks and designs will not sufficiently change urban, rural, or suburban schools to close achievement gaps ... in spite of significant evidence to the contrary, many educators continue to seek the single program … that will improve all students' achievement" (p. 19). Systemic change within and beyond education policy is necessary to ensure equitable access to opportunity in this country. Elements that must be addressed include equitable school funding, classroom-level desegregation, high expectations for all, rigorous curricula for all, supplemental help for students lacking in foundational skills, certified teachers for all, equitable resources for all, consistent standards for all, expanded access to learning opportunities outside of school hours, elimination of inequitable tracking, consistent and fair student discipline, and social and economic policies that address inequities outside of schools (Alemán, 2006; Haycock, 2001; Manning & Kovach, 2002; Mickelson, 2001; Reardon, 2016; Rothstein, 2004). Lastly, Milner (2010) has promoted paradigm shifts within education leaders to inspire greater motivation regarding and better approaches toward remedying opportunity gaps: "(1) rejection of color blindness; (2) ability and skill to understand, work through, and transcend cultural conflicts; (3) ability to understand how meritocracy operates; (4) ability to recognize and shift low expectations and deficit mind-sets; and (5) rejection of context-neutral mind-sets and practices" (pp. 13-14). Stripping away the entrenched opportunity gaps within the U.S. public education system will be no easy work, but it is necessary work if we are to fulfill the most basic promise of the institution. In their best form, magnet programs offer potential to contribute to this mission. However, this remains heavily under-researched, a fact I address in the "Gaps Specific to this Topic" section.

Opportunity-to-Learn Theoretical Framework

Recognizing that the opportunity gap represents a substantial threat to the success of students of color, I will examine the efficacy of magnet programs and their various themes through the lens of the opportunity-to-learn (OTL) framework. The OTL framework enables a researcher to "empirically connect context (such as teaching pedagogy and tracking) with learning outcomes to assess educational inequality within and between schools" (Griffin et al., 2007, p. 233). In this section, I will outline the roots, major tenets, significant works, and relevance of OTL as well as explicitly addressing how it informs this research. Considered together, these elements make the case for how this framework aligns with magnet research.

Roots of OTL

The concept of OTL originated in the 1960s through the work of John Carroll, an educational psychologist who developed a two-part model for learning based on teacher quality and opportunity to learn (operationalized as "time-on-task") (Jaffar, 2006; Scherff & Piazza, 2009). Additional researchers later expanded the construct to incorporate alignment between academic content taught and content tested, as well as other educational "indicators" (Jaffar, 2006; Scherff & Piazza, 2009). From there, OTL flourished as a long-standing framework for studies of international education (Dowd et al., 2014; Guiton & Oakes, 1995; Stevens, 1993; Wang, 1998). The International Association of Educational Achievement famously studied time-on-task using the OTL theory in the early 1960s, and the 1990 global school enrollment initiative Education for All emphasized its concepts as well (Dowd et al., 2014; Wang, 1998). By quantifying opportunities provided by various educational entities, the OTL framework enables comparisons between different education offerings, and it has done so for over 55 years (Guiton & Oakes, 1995).

American Political History of OTL. Within United States public education, OTL played a role in reform trends catalyzed by the infamous 1983 report *A Nation at Risk* (Guiton & Oakes, 1995; Herman & Klein, 1997; Jaffar, 2006; Stein, 2000; Stevens, 1993, 1996; Wiley & Yoon, 1995). In 1991, the National Council on Education Standards and Testing assembled an Assessment Task Force to explore accountability measures that might be incorporated in the burgeoning practice of standards-based student outcome accountability (Darling-Hammond, 1997; NCES, 1991). These measures were to be comprehensible to all stakeholders, reliable, valid, not burdensome on educators, minimally expensive, comparable over time, reflective of the complexity of schooling, and contextualized (NCES, 1991). In their 1991 report, the task force recommended six indicator categories to be included in President George H.W. Bush's America 2000 initiative, four of which related to opportunity to learn: quality of educational institutions, school readiness, societal support for learning, and educational equity (NCES, 1991). In the end, America 2000 developed six goals centering on enhancing public students' opportunities to learn but did not require any of the measures recommended by the task force (Darling-Hammond, 1997).

During the following administration, President Bill Clinton and Congress adapted and rebranded America 2000 as a new initiative entitled Goals 2000: Educate America

Act (Guiton & Oakes, 1995; Ladson-Billings, 2010; Porter, 1995; Wang, 1998). Goals 2000 proposed the explicit inclusion of OTL standards - conceptualized as resources, practices, and conditions needed for all to learn - in recognition of the fact that all schools needed certain minimum resources and systems for all students to meet the act's goals (Ladson-Billings, 2010; Porter, 1995; Wang, 1998). In effect, these standards would hold schools accountable for educational inputs (Guiton & Oakes, 1995). The inclusion of OTL standards became a major source of debate, regarded as too costly by Republican members of Congress (Ladson-Billings, 2010; Porter, 1995). As a compromise, Goals 2000 required states to adopt OTL standards if they wished to receive its funding but were not mandated to assess them (Guiton & Oakes, 1995; Porter, 1995). In effect, this tradeoff removed the teeth from this chance to hold states accountable for addressing opportunity gaps, and the concept of OTL standards was never actively pursued again.

Major Tenets of OTL

The basic premise of the OTL framework within the context of education research is to measure how educational inputs translate to student learning (Dowd et al., 2014; Griffin et al., 2007; Guiton & Oakes, 1995). Historically, the four key tenets of OTL research have been content coverage, content exposure, content emphasis, and quality of instructional delivery (Stevens, 1993, 1996; Wang, 1998). Individual researchers have at times proposed or implemented their own factors (Herman et al., 1996; Herman & Klein, 1997; Stevens, 1993; Wiley & Yoon, 1995). Despite any individual modifications, the fundamental aspect of the OTL framework is that it outlines "contextual factors that may effect [*sic*] student achievement, their 'opportunity to learn'" (Griffin et al., 2007, p. 233).

Significant OTL Works

Three research studies represent the foundation of modern OTL theory in the United States: Oakes's 1986 tracking study, the 1987 Los Angeles Unified School District's (LAUSD) Ten Schools Program, and a national schooling evaluation conducted by the NCES in the early 1990s (Jaffar, 2006; Oakes, 1986; Stevens, 1993, 1996). First, in her analysis of school tracking systems, Oakes (1986) found that tracking led to decreased opportunities to learn, which she operationalized similarly to the four tenets described above. The following year, LAUSD conducted its Ten Schools Program, in which the district analyzed its ten lowest-performing elementary schools (Stevens, 1993, 1996). In this project, the researchers identified what teachers were doing and how their practices could be changed to ensure higher quality instruction for all students (Stevens, 1993, 1996). Finally, Stevens et al. surveyed the research directors of 91 school districts, conducted a meta-analysis of the existing literature, and gathered data from the Third International Math and Science Study as members of the NCES Assessment Task Force (Jaffar, 2006; NCES, 1991; Stevens, 1996). Through this process, the research team identified four key variables impacting students' opportunities to learn, recognized as the four key tenets of the OTL framework today (Jaffar, 2006; NCES, 1991; Stevens, 1996). Considered together, these three studies constitute a precedent for OTL research, modeling how researchers can and should focus on educational inputs in order to identify avenues for improving American education.

Relevance of OTL

The literature has demonstrated that inequities are pervasive within public education in the United States, and these opportunity gaps have severe ramifications for students of color in particular (Aud et al., 2011; Harris, 2015; Mickelson, 2001; Orfield & Lee, 2007; Swanson, 2004). Measuring educational inputs redirects the focus from summative student assessment data to modifiable factors directly within educators' locus of control. As stated in the NCES report *Education Counts* (1991), "in counting the outcomes of education, it is easy to reverse ends and means. The intention is benign ... but the unintended effect is deceptive: We begin to value only what we *can* measure" (p. 5). In order to impact educational outputs, inputs must be changed, and such change is unlikely in the absence of a conscious effort to measure academic inputs. This drive is particularly relevant in the current and (presumably enduring) era of standards-based accountability (Jaffar, 2006). So, by evaluating educational inputs, researchers can draw focus to manners in which the system can be changed. Measuring educational equity through an OTL lens offers this potential.

How OTL Informs This Research

OTL measures are highly formative, clarifying how districts and schools are and are not addressing opportunity gaps, thereby revealing avenues for equitable change (Herman & Klein, 1997). In particular, the OTL framework empowers researchers to explore three primary questions: 1) Are content and instruction changing to align with educational expectations? 2) Are all students receiving opportunities that can prepare them to meet these new expectations? 3) Given this information, what changes in curriculum, material resources, and human resources should a school make? (Herman & Klein, 1997). This research strives to address the second of these questions within the context of magnet programs and their themes. Although magnet programs represent a potential opportunity for all students to meet educational expectations, I have found no research that proves this in terms of within-school racial achievement gaps. By examining the potential disparities in White, African American, and Hispanic students' educational outcomes within magnet programs, compared to within nonmagnet programs, I strive to answer the second OTL framework question within this vital context.

Effects of Magnet Programs on Student Achievement

At this point, the literature has demonstrated (1) that the history of magnets is intertwined with the inconstant process of school desegregation, (2) that magnet programs have had mixed effects on school desegregation, (3) that magnet programs have become more focused on academic enrichment than desegregation over time, and (4) that opportunity gaps persist in denying enriched academic opportunities to students of color. With this in mind, the final aspect of this literature review examines the effects that magnets have had on student achievement. Magnet programs' primary mission today appears to be academic innovation and achievement (Fleming, 2012; Frankenberg et al., 2008; Frankenberg & Le, 2008; Frankenberg & Siegel-Hawley, 2010; Goldring & Smrekar, 2000; Orfield & Lee, 2007; Rhea & Regan, 2007; Rossell, 2003; Siegel-Hawley & Frankenberg, 2012, 2013; Smrekar & Goldring, 2011; Wang et al., 2014; Yang et al., 2005), so it is worthwhile to examine the degree to which this promise has been realized.

This portion of the literature review includes 10 studies utilizing national datasets and 25 local evaluations of a district's or state's magnet programs, all found to have sufficiently rigorous methodology to give credence to their findings. Taken together, this research does not offer a consensus as to the impact of magnets on students' academic experiences (Ballou, 2007; Christenson et al., 2003). Among the 35 studies included, 17 suggest that magnet programs have positive effects on student achievement, two suggest negative effects, and 16 demonstrate mixed or neutral effects. These studies span from 1983 to 2017, with the bulk published in the late 1990s and 2000s. Although more recent data are preferable for generalizability to the current state of affairs, the height of interest in magnet program research in the United States took place in the 1980s and 1990s. The primary conclusion that can be drawn by aggregating the following information is that “in at least some times and places, students benefit from enrolling in magnet schools” (Ballou, 2007, p. 31).

Positive Effects of Magnet Programs on Student Achievement

In examining the research, 17 rigorous studies offered results indicating that magnet programs benefit students’ academic experiences. Four used national datasets and 13 evaluated magnets at a local level. They are detailed below.

Findings from National Analyses. Despite its age, the U.S. Department of Education’s report on 45 representative urban magnet programs funded by ESAA is regarded as a seminal work in the field (Blank et al., 1983). Blank et al. (1983) engaged in a mixed methods analysis of these magnets (selected from the 1,019 ESAA-funded magnet programs using stratified random sampling) and found higher test scores as well as preferable attendance, dropout, suspension, and transfer rates than district averages. Six years later, Blank (1989) followed up on these magnets and analyzed 12 additional districts’ magnet program evaluations, aggregating the data into a substantial meta-analysis. This research determined that the magnets studied produced higher achievement scores than their nonmagnet counterparts, even when controlling for students’ demographic data and prior academic achievement (Blank, 1989). Considered together, these two studies represent foundational research that justified magnet programs as a continuing venture in academic enrichment.

Two additional studies analyzed magnet programs' academic outcomes during the MSAP era. First, Gamoran (1996) compared the achievement data of 3,540 tenth grade students at urban high schools, 48 of which offered magnet programs and 308 which did not. The comparison of students' National Education Longitudinal Study (NELS) test scores identified a strong correlation between magnet program enrollment and increased reading and social studies scores (Gamoran, 1996). In 2012, Siegel-Hawley and Frankenberg (2012) surveyed leaders from 54 magnet programs receiving MSAP funding and found that 83% reported rising academic achievement. Participants credited this progress to the MSAP grant, particularly among programs that had more time to implement the funds (Siegel-Hawley & Frankenberg, 2012). Although these findings are limited, it appears that supplemental resources may enable greater student achievement at magnet programs.

Findings from Local Evaluations. Evaluations of individual states, districts, and schools by 13 researchers or research teams indicated that magnet programs produced positive effects on student achievement. In this section, these studies are organized by school level.

Elementary School Magnet Research Findings. Two studies of elementary magnet programs and two of hybrid elementary/middle school magnet programs revealed positive impacts of magnets on young learners. Larson and Allen (1988) compared the academic growth made by third through sixth graders attending Montgomery County (Maryland) Public Schools' 14 elementary magnet programs to their counterparts at nonmagnet schools. The math and reading achievement scores of magnet students increased more than those of the nonmagnet students, with gains comparable across racial groups (Larson & Allen, 1988). In a single-site analysis of a rural Louisiana elementary school, Haines and Kilpatrick (2007) found a remarkable turnaround effect following the adoption of an environmental science magnet program. The school's IOWA test scores (standardized to a 1-99 scale) increased from 26 points below the state average to 23 points above it, a transformation school leaders attribute to an increase in resources, community involvement, and student buy-in (Haines & Kilpatrick, 2007). In an effort to revitalize the district, Montclair Public Schools in New Jersey implemented a magnet program at every elementary and middle school (Clewell & Joy, 1990). Following this, both reading and math achievement scores increased and percentages of students below grade level decreased at every grade level tested and among all racial groups (Clewell & Joy, 1990). Test score analyses at six elementary or hybrid elementary/middle magnet programs in New Haven, Connecticut demonstrated an increase in all students' reading abilities and most students' math abilities (U.S. Department of Education, 2008). As a whole, these findings indicate that magnet programs may benefit young learners as they acquire fundamental reading and math skills.

Middle School Magnet Research Findings. Researchers have also found positive results at middle schools, according to the following three studies, as well as in one evaluation of middle and high school magnets together. Ballou (2007), Bifulco et al. (2009), and Martinez et al. (1996) implemented what is regarded as the most rigorous methodology in magnet research: comparing the outcomes of students who attended lottery-based magnets to those of students who applied to but lost the lottery and thus attended nonmagnet schools (Ballou, 2007). Lottery research "provides a way of disentangling the effect of the chosen school from the influence of factors that led to that choice … this makes unsuccessful participants a natural 'control group'" (Ballou, 2007, p. 1). Ballou (2007) utilized this approach and determined that attending one of five lottery-based middle school magnets resulted in increased math achievement at each magnet within the large school district studied. Bifulco et al. (2009) employed the same technique in studying Connecticut's interdistrict middle school magnets and found positive effects on reading achievement, although it is unclear if this resulted directly from magnet programming or from the less segregated environment the magnets offered. Further, Martinez et al. (1996) employed this methodology when studying a district offering two academically selective program-within-school middle school magnets in San Antonio, Texas. The researchers identified a correlation between attending one of these magnets with higher math and reading test scores, controlling for students' demographics and past academic achievement (Martinez et al., 1996). Martinez et al. (1996) also noted an anecdotal academic "pull-up" effect on nonmagnet students in the same buildings as the magnet students. All students attended classes together in subjects not associated with the magnets' foreign language and culture theme, suggesting that all students may have experienced more rigorous instruction and higher expectations in these shared courses (Martinez et al., 1996). Finally, Blank (1989) evaluated the six secondary magnet programs offered in Rochester, New York and found a below-average dropout rate at the magnets. Middle school magnet programs may offer more advanced educational opportunities than elementary magnets, as most students have already mastered the academic basics. This may represent one factor contributing to the strong results in these studies.

High School Magnet Research Findings. Magnet high schools may offer even more rigorous and specialized coursework. High school achievement research extends beyond test scores to include graduation and college access data, which may serve as one reason why a great deal of research occurs at this level. Eleven studies are examined here.

First, Bifulco et al. (2009) implemented lottery-based methods to explore a Connecticut high school magnet program. They found that students who attended the magnet outscored their peers who lost the lottery in both math and reading (Bifulco et al., 2009). In an analysis of two high school magnets in Austin, Texas, Blank (1989) identified higher ninth through eleventh grade math and science scores among magnet students than their nonmagnet counterparts. Given these findings, magnet programs appear to make a positive impact on high school test scores.

In addition to test score evidence, five studies demonstrated positive effects of high school magnets on graduation and dropout rates. Research in St. Louis and Los Angeles has shown increased graduation rates at magnet programs (Grooms & Williams, 2015; Silver et al., 2008; Yu & Taylor, 1997). For example, Silver et al. (2008) found that the graduation rate at Los Angeles Unified School District’s (LAUSD) nonmagnet high schools was 45%, compared to 73% at magnet programs. These studies show particularly strong benefits for Black students (Grooms & Williams, 2015; Yu & Taylor, 1997). Correspondingly, research in Cincinnati, St. Louis, Nashville, New York, and the Milwaukee metroplex found lower dropout rates at most magnet high schools compared to the districts’ nonmagnet high schools (Blank, 1989; Witte & Walsh, 1990; Yu & Taylor, 1997). In the Milwaukee area, in fact, large, urban magnet high schools had equivalent dropout rates to wealthy, White, suburban nonmagnet high schools (Witte &

Walsh, 1990). Given these findings, magnet programs may encourage students to stay in school and graduate.

Additionally, college information, preparation, access, and interest appeared more prevalent at high school magnet programs in Los Angeles and Portland, Oregon in research by Griffin et al. (2007) and Morales (2010). Griffin et al. (2007) studied juniors and seniors at one magnet and one nonmagnet high school in LAUSD. They concluded that, compared to their nonmagnet peers, those who attended magnet programs interacted with more college counselors, met more college representatives, were offered more Advanced Placement courses, and knew to seek out additional college information (Griffin et al., 2007). California sets specific prerequisites for admission to the University of California and California State University systems beyond the minimum requirements for a high school diploma (Griffin et al., 2007). Although 100% of magnet seniors and 90% of nonmagnet seniors studied graduated, 100% of magnet seniors completed the UC/CSU admission requirements compared to a mere 8% of seniors at the nonmagnet high school (Griffin et al., 2007). Additionally, magnet programs may encourage students to seek out more rigorous college programs and pursue more competitive careers, according to an analysis of students at a science and technology high school magnet in Portland (Morales, 2010). These students reported increased interest in and preparation for STEM careers due to the magnet's required internship, problem-solving focus, and self-directed projects (Morales, 2010). So, not only does research indicate that magnet programs benefit students while they are in high school, but this experience appears to advance their post-high school trajectories.

K-12 Magnet Research Findings. Finally, researchers documented the positive effects of magnet programs in evaluations of six school districts. First, students at the 83 magnet programs in the Houston Independent School District outperformed district averages on both state standardized tests and grade-level equivalency exams, with greater differences appearing in later grades (Stanley, 1989). These magnet students had standardized test score averages seven to 23 percentage points (depending on the grade level) above the district average and scored 0.3 to 4.9 grade-level equivalents (depending on the grade level) above the district average (Stanley, 1989). Similarly, Witte and Walsh (1990) compared test scores of magnet and nonmagnet students at 204 schools in the Milwaukee metroplex and found “an undeniable and dramatic difference” (p. 205). Magnet students scored .44 to .77 standard deviations higher in reading and .81 to 1.15 standard deviations higher in math than their nonmagnet peers, controlling for student characteristics and organizational variables, such as parental involvement and teacher control (Witte & Walsh, 1990). In addition, Yu and Taylor (1997) and Poppell and Hague (2001) found that low-income magnet students outscored their nonmagnet peers in Cincinnati, St. Louis, Nashville, and Jacksonville, Florida. Grooms and Williams (2015) determined that Black students in particular benefited from St. Louis’s high school magnets. Lastly, Grotto (2002) analyzed arts magnets schools of all levels around the country and identified increased test scores, college and career readiness, and social/emotional benefits. Taken together, the preceding studies indicate that magnet programs have the potential to benefit students on a variety of outcomes.

Negative Effects of Magnet Programs on Student Achievement

A minority of research findings suggest explicitly negative effects of magnet programs on student learning. One national study and one district evaluation are of sufficient rigor to be included here.

Findings from a National Analysis. Walton and Ford (2014) analyzed student achievement data from the 2010 MSAP grantee cohort. Although this report included mixed findings, the primary result was negative (Walton & Ford, 2014). The authors found a decrease in the average percentages of students who met or exceeded grade-level standards by 1 to 10% in both reading and math over the three-year grant period (Walton & Ford, 2014). However, the authors noted a number of possible confounding factors. These included state-level changes in accountability protocols, new benchmarks, and teacher retraining in pedagogy and content to align with their magnet theme (Walton & Ford, 2014). Although Walton and Ford (2014) indicated a negative impact of magnets on a substantial number of schools, these limitations impede generalizability.

Findings from a Local Evaluation. In addition, one district-level analysis produced negative results. Adcock and Phillips (2000) evaluated academic outcomes from the 28 elementary schools with magnet programs and the 89 without in Prince George's County Public Schools in Maryland. Controlling for first grade IQ, the researchers found that the district's nonmagnet students outperformed its magnet students on state standardized tests (Adcock & Phillips, 2000). However, Adcock and Phillips' (2000) use of hierarchical linear modeling has since become criticized by peers within the magnet research community due to arguably inconsistent independent variable operationalization (Ballou, 2009). In short, both of the studies indicating that magnet programs negatively impact student achievement should be considered with caution.

Mixed and Neutral Effects of Magnet Programs on Student Achievement

To echo the earlier quote from Ballou (2007), the following research reflects how "in at least some times and places, students benefit from enrolling in magnet schools" (p. 31). In these 16 studies - five with national datasets and 11 at a more local level - the authors identified mixed or neutral effects of magnet programs on student achievement.

Findings from National Analyses. Five national studies, presented chronologically here, have suggested that MSAP-funded magnets may not yield a single, clear-cut effect on student achievement. Blank et al. (1983) conducted the first national analysis of magnet programs on behalf of the U.S. Department of Education. In this mixed methods study, they evaluated 45 magnets representing 15 urban school districts (Blank et al., 1983). 19% of the programs reported reading standardized test scores below their district average and 22% did so for math (Blank et al., 1983). Many of the remaining 81% and 78% of programs, respectively, reported above-average test scores in reading and math, a discrepancy correlated with magnet theme coherence, school principal quality, and degree of special treatment afforded by the district (Blank et al., 1983). Two decades later, Christenson et al. (2003) published a similar evaluation of the 1998-2001 MSAP grantee cohort and the cumulative 292 magnet programs they supported for the Department of Education. At this time, grantees were required to set specific targets for language arts and math achievement scores for each school, and Christenson et al. (2003) determined that 51% of magnets met most of their language arts goals while 36% met most of their math goals. This growth did not meet the statistical threshold to be considered significantly different from the performance of nonmagnet students once researchers controlled for student demographic data, suggesting a neutral effect (Christenson et al., 2003). Similar findings emerged in an analysis of 21 elementary magnet programs from the 2004-2007 and 2007-2010 MSAP cohorts (Betts et al., 2015). Specifically, the researchers determined that math achievement scores did not change, and language arts scores only grew at magnets sited at "previously disadvantaged" schools (Betts et al., 2015). Finally, Wang et al. (2014, 2017) published two evaluations of 24 magnet programs funded through the 2010-2013 MSAP grant cycle, one at the conclusion of the first grant year and the second after the third year. Both studies revealed such heterogeneity among test scores, particularly among African-American and low-income students, that the researchers concluded an overall finding of no effect on student achievement (Wang et al., 2014, 2017). Together, these large-scale studies of MSAP recipients support the notion that magnet programs may vary to such a high degree as to prevent a universal conclusion of their efficacy.

Findings from Local Evaluations. Eleven rigorous local-level evaluations produced mixed or neutral findings as well. These are organized by school level.

Elementary School Magnet Research Findings. In analyses of elementary magnet programs in the Wake County Public School System, across Maryland, and in the San Diego Unified School District, researchers identified neutral effects (Betts et al., 2006; Penta, 2001; Yang et al., 2005). Penta (2001) compared achievement data across Wake County PSS's 30 elementary schools with magnets and 44 elementary schools without, controlling for student demographic data. She identified no differences in reading or math outcomes, which she presented in a positive light, stating that "students at program magnets are able to benefit from these unique offerings [e.g., electives, supplemental classes/curriculum related to magnet theme] and still sustain their academic achievement in core areas" (Penta, 2001, p. 9). In evaluating magnets in the San Diego USD, Betts et al. (2006) found no impact on math or reading scores as well. In a more limited study, Yang et al. (2005) compared fifth grade achievement at seven elementary magnet programs across Maryland with their nonmagnet counterparts. Their results showed that six of the programs had neutral to slightly negative effects on reading and math achievement (Yang et al., 2005). These three sets of findings suggest that magnet programs at the elementary level are far from guaranteed to impact student achievement.

Middle School Magnet Research Findings. In two studies of middle school magnet programs, researchers discovered neutral effects on student learning. Ballou et al. (2006) analyzed four middle school magnets. Once the researchers controlled for student demographic variables, the magnets offered no statistically significant benefits (Ballou et al., 2006). Additionally, in their comprehensive study of the San Diego USD's magnet programs, Betts et al. (2006) found that the district's middle school magnet programs had no impact on math or reading achievement. This research echoes the conclusion of the elementary magnet findings above.

High School Magnet Research Findings. Additionally, researchers have demonstrated mixed or neutral effects of high school magnets on students' test scores, graduation rates, and post-education outcomes in seven rigorous district-level studies. A subset of these have found no impact on reading achievement (Betts et al., 2006; Cullen et al., 2005; Goldschmidt & Martinez-Fernandez, 2004; Kemple & Snipes, 2000) and minimal to no impact on math achievement (Betts et al., 2006; Cullen et al., 2005; Kemple & Snipes, 2000). It is worth noting that Betts et al. (2006), Cullen et al. (2005), and Kemple and Snipes (2000) utilized the highly rigorous lottery-based methodology described previously. Interestingly, when Crain et al. (1992) analyzed data from 4,258 ninth grade students who participated in a lottery for New York City's career academy magnet programs, they revealed an increase in academic achievement but a widening of gaps between students based on reading ability. Magnet students' math and reading growth substantially outpaced that of their nonmagnet counterparts, but the gap between magnet students with below-average and average general reading abilities widened significantly, a change that did not occur at the nonmagnet high schools studied (Crain et al., 1992). Further, Cullen et al. (2005) conducted a lottery-based evaluation of Chicago Public Schools' high school magnet programs and found that magnet students were no different from students who lost the lottery and subsequently attended nonmagnet schools in terms of attendance, credit accumulation, or course-taking patterns. In contrast, Kemple and Snipes (2000) determined that magnet lottery applicants who won admission to New York City Public Schools' career academy magnets exhibited better attendance and course completion than their peers who did not obtain a seat in a magnet program. As with the research regarding elementary and middle school magnet programs described in the previous two sections, these findings indicate that academic benefits are far from guaranteed with magnets and that such effects are likely context-specific.

Similar to the findings of mixed or no effect on test scores, in two rigorous lottery-based evaluations, Crain et al. (1992) and Cullen et al. (2005) reported that high school magnets had little to no effect on graduation or dropout rates in New York Public Schools and Chicago Public Schools. In the Chicago context, graduation rate had no correlation with attending a lottery-based magnet program, although the research team found that magnet attendance was related to decreased school discipline and arrest rates (Cullen et al., 2005). Students who attended a New York City career academy magnet following a lottery win earned more graduation credits and were less likely to drop out

(Crain et al., 1992). However, the decrease in dropout rate was limited to 6% (Crain et al., 1992). Unlike research presented in the "Positive Effects of Magnet Programs on Student Achievement" section of this chapter, these studies suggest that magnet students may be no more likely to complete high school than their peers.

Finally, researchers have explored post-education outcomes among magnet and nonmagnet high school graduates in two rigorous follow-up studies. Bank and Spencer (1997) surveyed 312 graduates from a campus offering a program-within-school magnet. They found that although students who had participated in the magnet program reported higher educational aspirations than their nonmagnet classmates, there was ultimately no difference in actual education attainment between the two groups (Bank & Spencer, 1997). Kemple and Scott-Clayton (2004) conducted a four-year follow-up study on Kemple and Snipes's 2000 career academy research. They found no difference in educational attainment or females' earnings when comparing students who did and did not win admission to a career academy magnet (Kemple & Scott-Clayton, 2004). However, male magnet graduates reported 18% higher earnings than their nonmagnet counterparts, and individuals who had been identified as at risk of dropping out when they entered high school reported 14-16% higher earnings if they attended the magnet (Kemple & Scott-Clayton, 2004). Taken together, these magnet high school studies suggest a need for further research to confirm whether magnet students and graduates experience short- and long-term academic benefits.

K-12 Magnet Research Findings. One additional study provides evidence regarding the academic effects of magnet programs. Rhea and Regan (2007) evaluated the 35 magnet programs at all school levels within Wake County Public Schools by comparing them to pair-matched schools without magnets. In this analysis, they identified almost entirely neutral outcomes in terms of students' test scores, students' academic growth, and the number of schools achieving Adequate Yearly Progress, although they discovered some slightly negative findings as well (Rhea & Regan, 2007). This final district evaluation echoes the results of the previous ten: there may be no consistent, significant effect of magnet programs on student learning, but rather, effects may vary substantially by within- and beyond-program variables.

School-Level Factors That May Impact Academic Opportunities and Student Achievement at Magnet Programs

The primary takeaway from the preceding section, "Effects of Magnet Programs on Student Achievement," should be the lack of a conclusive, universal magnet program effect. This section presents research regarding a variety of school-level factors that may impact the degree to which magnet programs address opportunity gaps and enhance student achievement. It appears that "the devil is in the details. It matters how the choice plan is designed. Different outcomes accompany different procedures" (Weiss, 1996, p. vii). Thus, understanding the impact of variables within magnets may reveal ways to strengthen their efficacy.

School Setting

School location appears relevant to magnet program development and implementation. Witte and Walsh (1990) described the U.S. public education system as offering "two very separate educational worlds - one in the city and one in the suburbs" (pp. 192-193). Recently, Wang and Herman (2017) showed that 55% of magnet programs are located within urban areas. This number is roughly consistent with Walton and Ford's

(2014) determination that 65% of 2010-2013 MSAP grantee magnets were sited at urban schools. The discrepancy between the two values should not inspire concern though, as MSAP cohorts are not meant to be representative samples of the larger population of American magnet schools. Much of the past and present of magnet schools can be found in the urban hearts of metroplexes, a truth that informs the sampling for this study, which will be detailed in Chapter Three.

Desegregation Plan Type

Considering the fact that magnet programs exist in the United States due to the school desegregation movement and, in many cases, individual programs have been developed due to desegregation plans, variables related to desegregation efforts are worth addressing when exploring the heterogeneity of magnet education. Magnets can be implemented through both mandatory, court-ordered desegregation plans and voluntary desegregation plans adopted of a school district's own volition. Comparing data from the 1990s and from the 2000s to 2010s proves somewhat ironic. Although each era prioritized a different plan type, the most prevalent plan format has typically been the less effective one of time. Mandatory plans were more widespread in the 1990s (Rossell, 2003), yet Steele and Eaton's (1996) analysis of Department of Education data proved that they were 7% less effective at achieving desegregation goals than voluntary plans during this time. Both aspects of this trend appear to have since reversed. 76% of 2007-2010 MSAP grantee districts and 77% of 2010-2013 MSAP recipients were under a voluntary plan (Walsh, 2007; Walton & Ford, 2014). Unfortunately, mandatory plans now appear to be most effective (Clotfelter, 2004; Walton & Ford, 2014). Walton and Ford's (2014) review of the 2010-2013 MSAP grantee cohort for the Department of

Education found that districts under mandatory plans were 15% more effective at meeting desegregation targets. They suggested that this was due to:

> "aggressive strategies, timelines, and strong district support … established policies and procedures … systems in place to reduce minority group isolation … regular compliance monitoring by an oversight agency … [and] immediate adjustments if their school enrollments [were] not moving in the right direction." (Walton & Ford, 2014, p. 37)

Whether these incongruities are labeled ironic or tragic, this research suggests that school districts and the courts have yet to enact methods that align with the needs of the time, and programmatic alignment with equity goals is certainly relevant in addressing the degree to which magnet programs realize their potential for mitigating opportunity gaps.

Magnet Program Format

Two primary magnet program formats exist: whole-school magnets (in which all students enrolled at the campus receive magnet programming) and program-within-school magnets (in which a subset of the campus's students receive magnet programming), commonly labeled PWS (Goldring, 2009). Among either type, district policy may permit the magnet to serve students who live within an assigned attendance zone, who apply into the program from across the district, or a combination of the two (Goldring, 2009). The two formats appear to differ in terms of both desegregation and academic outcomes.

According to survey data as well as analyses of district enrollment patterns in St. Louis, Cincinnati, and Prince George's County, Maryland, whole-school magnets appear to be more effective desegregation tools than PWS magnets (Frankenberg et al., 2008;

Goldring, 2009; Rossell, 2003; Smrekar & Goldring, 2000). Contrary results were discovered in survey data by Siegel-Hawley and Frankenberg (2013), however, who found that PWS magnets were less likely to be racially isolated. In a unique element of their report for the U.S. Department of Education, Walton and Ford (2014) compared the desegregation efficacy of magnet programs with attendance zones and those without, who required applications from each student. Magnets without attendance zones were more effective at achieving desegregation targets, perhaps because they were able to circumvent residential segregation patterns (Walton & Ford, 2014). Finally, it is worth noting that the preceding studies analyzed school-level desegregation. Second-generation may underlie these findings, particularly within PWS magnets (*People Who Care v. Rockford Board of Education School District No. 205*, 2001; West, 1994).

Despite their poor desegregation outcomes, PWS magnets appear to experience stronger academic outcomes than whole-school magnets. Although the majority of magnet benefits may be sequestered to students participating in the magnet, well-implemented PWS magnets may offer a “pull-up” effect for students not enrolled in the campus’s magnet program (Goldschmidt & Martinez-Fernandez, 2004; Martinez et al., 1996). Goldschmidt and Martinez-Fernandez (2004) and Martinez et al. (1996) found this positive effect among PWS magnets in California and San Antonio, where nonmagnet students attained greater academic achievement, likely due to classes which nonmagnet and magnet students attended together (Goldschmidt & Martinez-Fernandez, 2004; Martinez et al., 1996). When designing new magnets, districts should select formats carefully, as they offer implications for which students are served and what academic opportunities are afforded them.

Magnet Program Theme

Magnet program themes may center on a curricular focus or a pedagogical approach. The selection of a theme is typically a local decision, ideally chosen with stakeholder input. Walton and Ford's (2014) Department of Education evaluation of the 2010-2013 MSAP cohort offers the most recent rigorous data on theme prevalence. Among this cohort, the following five themes were most common: science, technology, engineering, and math (38%); arts and humanities (22%); International Baccalaureate (17%); career and technical (7%); and foreign language and cultural studies (5%) (Walton & Ford, 2014). Walton and Ford (2014) also determined that career and technical magnets were most effective in desegregating and improving language arts and math performance. Their data appear roughly consistent with statistics recently published by the Magnet Schools of America organization, which identified STEM (30%), visual and performing arts (16%), International Baccalaureate (12%), gifted and talented (8%), and foreign language (7%) as the most prevalent magnet themes in the United States (Magnet Schools of America, 2017).

To attract and retain students, themes should be chosen carefully and implemented with fidelity. A former superintendent recommended that magnets "scream the theme," crediting his district's magnet success to "keeping the focus of the theme alive … once the theme of the school is lost, the magnet concept dies" (Jackson, 2007, pp. 34-35). A magnet in name only is unlikely to attain the same outcomes as a magnet committed to fulfilling its promise of enriching academic opportunities.

Magnet Program Age

Initial research indicates that duration of implementation also impacts the efficacy of magnet programs. At the outset, certain magnets are likely to experience little to no change, as meaningful implementation requires time to retrain teachers and recruit students (Betts et al., 2015; Morrison, 1996). Morrison (1996) first demonstrated this phenomenon in Missouri's Kansas City School District, where magnet age explained a good deal of the variation in desegregation efficacy. In their 2014 analysis, Walton and Ford compared new campuses with magnets, established campuses with new magnets, and established magnet campuses. Of the 2010-2013 MSAP magnets, new campuses opening with magnets were more successful in fostering desegregation than both previously nonmagnet schools adopting magnets and existing magnets updating their programs (Walton & Ford, 2014). When possible, magnet researchers should consider magnet age in their analyses, as this variable appears to have implications for which students have access to magnets' potentially enriching academic opportunities.

Leadership

School leaders greatly impact learning environments, including magnet programs. Leadership choices can dramatically shape students' learning opportunities and their subsequent outcomes. Research exploring magnet program leadership was quite popular in the 1980s and 1990s, and two major findings can be taken away from this research. First, effective magnet leaders empower teachers through autonomy in decision-making and distributed leadership (Bauch & Goldring, 1996; Chubb & Moe, 1990; Dentler, 1990; Hausman et al., 1997; Smrekar & Goldring, 1999; Wehlage & Smith, 1992). Second, successful magnet programs tend to experience a degree of latitude from their district headquarters so that principals can be more autonomous as well (Blank et al., 1996; Clewell & Joy, 1990; Wehlage & Smith, 1992). Strong leadership through these two avenues has correlated with both desegregation and academic benefits for magnet programs (Blank et al., 1983; Dentler, 1990).

However, few studies have explored magnet leaders in the new millennium, and researchers have found that leadership may be waning within twenty-first century magnet programs. In a rigorous analysis of schools in St. Louis and Cincinnati, Hausman (2000) discovered virtually no differences between magnet and nonmagnet principals in terms of instructional leadership, collaborative practices, parent empowerment, central office interaction, or school recruitment. Hausman and Goldring (2001) found that magnet teachers in St. Louis and Cincinnati rated their principals as less effective than their nonmagnet peers. Magnet teachers who rated their principals as effective described their programs as “exemplified by clear goals for student achievement, teacher talk about instructional objectives, clear guidelines about what teachers are to emphasize in their teaching, and no burdensome paperwork” (Hausman & Goldring, 2001, p. 415). However, contemporary magnet school administrators appear to be experiencing less autonomy (Frankenberg et al., 2008), turning over at higher rates (Walton & Ford, 2014), and focusing most on state standardized assessments (Christenson et al., 2003). These factors may be contributing to the apparent weakening of magnet program leadership, which may hinder program efficacy in terms of academic and desegregation benefits.

Instructional Practices

Due to their unique themes, magnet programs often offer instruction that differs from traditional teaching. The most recent large-scale test of this hypothesis was in Christenson et al.’s (2003) evaluation of the 1998-2001 MSAP cohort for the Department of Education. They determined that MSAP teachers solicited more higher-order thinking, used more technology to engage students, and implemented more varied assessments than nonmagnet teachers (Christenson et al., 2003). Additionally, they found that grantee programs had smaller student-to-teacher ratios than nonmagnet schools in their districts (Christenson et al., 2003). Earlier research identified similar trends (Blank et al., 1996; Blank & Archbald, 1992; Gamoran, 1996; Smrekar & Goldring, 1999; Stanley, 1989).

Additionally, researchers have explored the degree to which educators implement their magnet theme with fidelity. In two studies of MSAP grantee magnets, Wang et al. (2014, 2017) determined that fidelity of theme implementation as well as the work of magnet coordinators collectively explained 60% of the variation in students' math performance and 40% of the variation in reading. If magnets are to offer enhanced academic opportunity compared to nonmagnet schools, magnet educators must truly adopt the theme and adapt their instruction.

Climate and Community

School climate encompasses "perceptions of a school's academic norms, expectations and beliefs" (Mattison & Aber, 2007, p. 2) as well as "the quality and consistency of interpersonal interactions within the school community" (Haynes et al., 1997, p. 322). In an extensive literature review of school climate research, Haynes et al. (1997) identified correlations between school climate and students' self-concepts, behavior, attendance rates, discipline rates, academic achievement, sense of safety, failure rates, perceptions of expectations, and mental health. School climate can certainly impact educational experiences, in short.

Magnet programs may lend themselves to developing a more positive school climate and a stronger sense of community. Christenson et al.'s (2003) Department of

Education evaluation of the 1998-2001 MSAP cohort identified strong school climate and professional community among MSAP magnets, especially at the elementary level, which strengthened over the grant period. Interestingly, the researchers determined that climate was the factor most correlated with reading and math growth within the 292 magnets analyzed (Christenson et al., 2003). Research teams have also found above-average levels of community at magnets in New York City and Nashville (Hausman & Goldring, 2000; Kemple & Snipes, 2000). Kemple and Snipes (2000) discovered higher expectations, more personal attention, more collaborative activities, and more engaged peers at a career academy PWS magnet in New York City than within the nonmagnet portion of the campus. These supports correlated with academic benefits (Kemple & Snipes, 2000). Positive correlations between magnets and school climate have also been identified in previous decades during the height of magnet research (Bauch & Goldring, 1995; Crain et al., 1992; Doyle & Levine, 1984; Gamoran, 1996; Larson & Allen, 1988; Wehlage & Smith, 1992; Yu & Taylor, 1997).

Additionally, parents and teachers alike have reported higher levels of satisfaction at magnets than at nonmagnet schools. Hausman and Goldring (2000) and Teske et al. (2007) surveyed parents of children attending magnet programs in large urban districts. In Hausman and Goldring’s (2000) study, 80% of parents rated their child’s magnet program an A or B, with greater satisfaction rates among parents who reported choosing the school based on academic and value reasons. In their survey of 800 low-income families participating in school choice, Teske et al. (2007) found that 88% of parents felt satisfied with their selection. Findings in earlier magnet research corresponded with these results as well (Blank, 1989; Clewell & Joy, 1990; Larson & Allen, 1988; Martinez et al., 1994, 1996).

Similarly, magnet teachers appear content with their programs' climates, especially teachers at integrated schools (Frankenberg et al., 2008). These results are also consistent with prior research (Blank, 1989; Chubb & Moe, 1990; Doyle & Levine, 1984; Smrekar & Goldring, 1999). One potential factor in teachers' heightened satisfaction may be that magnet programs tend to have lower rates of student disciplinary infractions (Blank, 1989; Doyle & Levine, 1984; Engberg et al., 2011; Morris & Goldring, 1999). Considered together, these school climate indicators suggest that magnets may be more positive environments in which to teach and learn than comparable nonmagnet schools.

Family and Community Involvement

A final factor that may impact magnet program outcomes is the degree to which students' family and community members are involved with schools. In this area, results within the extant literature are highly mixed. A number of researchers have found that families that choose a school, such as one with a magnet program, may feel more invested and subsequently seek out more involvement opportunities (Blank et al., 1983; Clewell & Joy, 1990; Hausman & Goldring, 2000; Martinez et al., 1996; Poppell & Hague, 2001; Yu & Taylor, 1997). This may result from a self-fulfilling prophecy of school climate, a desire for confirmation that the school was a good choice, or an effort "to ensure that the school remains consistent and congruent with their values" (Hausman & Goldring, 2000, p. 110). In addition, some researchers have suggested that a school's community may become more involved if the school offers a magnet (Blank et al., 1983; Haines & Kilpatrick, 2007; Poppell & Hague, 2001). Magnet programs may inherently foster family and community involvement through these avenues.

However, several research teams have quantified low levels of involvement at magnet programs, typically in districts where nonmagnet schools also experience little family and community involvement (Bauch & Goldring, 1996; Christenson et al., 2003; Hausman & Goldring, 2000, p. 200; Smrekar & Goldring, 1999). Through interviews, Smrekar and Goldring (1999) found that a number of parents regarded selecting a magnet program as handing off responsibility for children's schooling. A second theory is that magnet families often enroll their children in schools further from their homes than their zoned school, breaking patterns of geographic community (Smrekar & Goldring, 1999). To reconcile these contrary findings, Bauch and Goldring (1995) suggested that "the opportunity to choose a school in itself may not support greater parental involvement; rather, what the school does to respond to parents may be the more important key" (p. 5). As with a great deal of the research presented in this chapter, although contemporary magnets share a history, their current realities vary greatly, and many nuances within the heterogeneity of magnet programs today hold implications for the academic opportunities they offer.

Gaps Within the Literature

General Concerns with Magnet Research

Before closing, one must acknowledge the limitations inherent in magnet program research. These include research age, sample size concerns, varied methodologies and measurement techniques, poorly controlled and possibly confounding variables, and improper selection of dependent variables.

First, the vast majority of magnet research was conducted between the 1970s and 1990s (Ballou, 2009; Wang et al., 2014). This timing corresponds with a major boom in magnet program implementation (Blank et al., 1983, 1996; Blank & Archbald, 1992; Smrekar & Goldring, 2000; Steele & Eaton, 1996), as researchers and practitioners alike were likely curious to explore the effects of this newly popular school choice option. In addition to potentially outdated findings, this older research may lack the rigor of more recent methodologies (Wang et al., 2014).

Second, much of the research consists of evaluations of individual districts' magnet programs, offering a quite limited sample size (Ballou, 2009; Betts et al., 2015). Additionally, both district-level and national magnet research tends to employ school-level data in lieu of student data (Ballou et al., 2006).

Third, magnet research has no universal standard for methodology or for operationalizing school diversity. Reardon and Owens (2014) characterized the variety of models, parameters, and outcomes as "theoretical confusion in the literature" (p. 214). Among desegregation research, researchers employ a number of indices to quantify diversity within a school's student body, including exposure, racial isolation, segregation, racial imbalance, dissimilarity, and entropy (Christenson et al., 2003; Clotfelter, 2004; Clotfelter et al., 2006; Frankenberg, 2008). This renders comparisons across studies challenging.

Fourth, magnet programs are likely impacted by so many variables (including but certainly not limited to the nine school-level factors detailed in the previous section) that it is immensely challenging to control for their effects and address downstream mediators (Ballou et al., 2006; Betts et al., 2015; Blank & Archbald, 1992; Reardon & Owens, 2014; Wang & Herman, 2017). Students' prior academic achievement in particular is rarely controlled for in a sufficiently rigorous manner (Ballou et al., 2006). In some sites, magnet and nonmagnet populations may also be sufficiently different in terms of background variables to compromise assumptions underlying statistical models (Yang et al., 2005).

Finally, even when a study is sufficiently recent, has a substantial sample size, and employs a rigorous methodology that controls variables well, its results may still be of little worth. The outcome variables often selected in magnet research may not align with the mission of numerous programs, as “many magnet school programs are not specifically aimed at building the skills reflected on standardized tests” (Blank & Archbald, 1992, p. 87). Siegel-Hawley and Frankenberg (2013) articulated this limitation eloquently:

> One of the ironic aspects of magnet school evaluation is that it is commonly limited to math and reading scores. If a student becomes fluent in another language, learns how to operate a commercial enterprise, develops a deep understanding of history or government, or learns to sensitively perform a role from Shakespeare, that counts for nothing in traditional evaluations. This means that virtually all appraisals of the educational effects of magnets ignore the very aspects that make the programs magnetic. (p. 136)

Additionally, changes in state standardized tests and increasing pressure to “teach to the test” also compromise the validity of test scores as a measure of academic enrichment, as does regression to the mean (Ballou et al., 2006; Harris, 2006; Rothstein, 2004).

All of these potential limitations must be kept in mind when examining the research literature.

Gaps Specific to This Topic

Despite the research exploring the academic effects of magnet programs detailed above, a substantial gap exists. A number of researchers (Betts et al., 2015; Bifulco et al., 2009; Gamoran, 1996; Wang et al., 2014, 2017; Wang & Herman, 2017) have expressed a need for future studies to investigate "specific [themes] of magnet schools instead of describing the average of this diverse category of schools" (Gamoran, 1996, p. 14). School districts around the United States employ a wide range of magnet themes, and it is unlikely that STEM, Montessori, career, arts, International Baccalaureate, foreign language, Paideia, and literature magnet programs provide equivalent academic opportunities and yield the same academic outcomes. With new findings that clarify what magnet themes mitigate opportunity gaps and enhance students' educational experiences, districts may become better equipped to fulfill the promise of magnet programs.

Chapter III

Methodology

In this chapter, I describe the methodology for this thesis. I begin with the research questions, research design, and hypotheses. Then, I offer descriptions of the setting and participants, followed by information regarding the dataset. Finally, I outline the variables, instrumentation, and procedure of this work. Considered together, these elements enable me to explore the relationship between magnet themes and within-school racial achievement gaps in urban public schools across Texas.

Research Questions

As stated in Chapter One, the research questions for this study are as follows:

Research question 1: What is the nature of the relationship between magnet status (i.e., whether a school offers a magnet program or not) and within-school racial achievement gaps, controlling for contextual variables known to influence academic outcomes?

Research question 2: Does this relationship vary across magnet program themes?

Opportunity gaps remain pervasive within American public education, resulting in disparities in academic achievement between White students and students of color (Alemán, 2006; Aud et al., 2011; Berends, 2014; Darling-Hammond, 1997; Gilliam, 2016; Gilliam et al., 2016; Gregory et al., 2006; Haycock, 2001; Kahlenberg, 2001; Ladson-Billings, 2010; Lee, 2002; Manning & Kovach, 2002; Mattison & Aber, 2007; Mickelson, 2001; Milner, 2010; Noguera & Wing, 2006; Oakes, 1986; Orfield et al., 2008; Orfield & Lee, 2007; Petras et al., 2011; Rothstein, 2004; Rubin et al., 2006; Scherff & Piazza, 2009; Skiba et al., 2002, 2011; U.S. Department of Education, Office of Civil Rights, 2014; Wang, 1998; Williams, 2002a, 2002b; Wing, 2006; Wu et al., 1982). Magnet programs have been advertised as potentially enhancing academic opportunities for students of color and other historically underserved students, but the degree to which magnets achieve this remains unclear. As detailed in Chapter Two, a great deal of heterogeneity exists within the extant literature on magnet programs and their academic effects. Much of this literature, additionally, measures average academic outcomes at the school level, with minimal exploration of racial achievement gaps within schools. For this reason, I ask the first research question, aiming to compare racial disparities in standardized test scores at urban Texas public schools with and without magnet programs. I control for variables known to impact academic performance to distill the relationship between magnet status and racial achievement gaps.

I pose the second research question with the aim of illuminating which magnet themes are associated with decreased racial academic gaps within schools. A wide variety of curricular and pedagogical themes exist within contemporary magnet programs. According to the Magnet Schools of America organization, the most prevalent magnet themes are STEM (30% of magnets), visual and performing arts (16%), International Baccalaureate (12%), gifted and talented (8%), and foreign language (7%) (Magnet Schools of America, 2017). Similarly, Walton and Ford (2014) found that the most common themes among the 151 MSAP grant recipients were STEM (38%), arts and humanities (22%), International Baccalaureate (17%), career and technical (7%), and foreign language and cultural studies (5%) in their analysis for the U.S. Department of Education (Walton & Ford, 2014). It is unlikely that each theme impacts students' academic opportunities and outcomes in the same manner. So, I seek to answer the second research question, with the long-term goal of offering guidance to school districts as they select magnet themes to implement in their schools to enhance students' learning opportunities.

Research Design

To measure the potential relationships outlined in the research questions, I conduct a cross-sectional, quantitative secondary analysis using a large dataset. Researchers and practitioners alike often measure students' academic outcomes quantitatively, with standardized test score data serving as an unofficial industry standard. For this work, I compile a dataset using demographic and achievement data from the Texas Education Agency's data warehouse. This digital warehouse includes the data used to determine school and district ratings, further indicating its significance in the context of Texas public education. In the "Dataset" section below, I detail the information I download from TEA, which encompasses over 30 values for each of the 1,357 schools studied. A quantitative approach facilitates a rigorous analysis of a dataset of this size, enabling a more comprehensive picture of racial achievement gaps within contemporary magnet programs. Such a design lends itself better to generalizability than a qualitative study or one with a smaller dataset, as well.

Hypotheses

For the first research question, the null hypothesis is that there is no significant relationship between magnet status and within-school racial gaps in state standardized test scores when controlling for other contextual variables known to influence academic outcomes. This can also be expressed as: H_0 = There is no significant relationship between magnet status and within-school racial gaps in state standardized test scores when controlling for contextual variables known to influence academic outcomes. The alternative hypothesis for this is that there is a significant relationship between magnet status and within-school racial gaps in state standardized test scores when controlling for contextual variables known to influence academic outcomes. As an equation, this can be written: H_1 = There is a significant relationship between magnet status and within-school racial gaps in state standardized test scores when controlling for contextual variables known to influence academic outcomes.

The null hypothesis for the second research question is that the relationship between magnet status and racial gap in state standardized test scores does not vary across magnet program themes. This can be represented as: H_0 = The relationship between magnet status and racial gap in state standardized test scores does not vary across magnet program themes. The alternative hypothesis for this question is that the aforementioned relationship does vary across magnet program themes. This can also be expressed as: H_1 = The relationship between magnet status and racial gap in state standardized test scores varies across magnet program themes.

Setting and Participants

Setting

Magnet programs were born in urban school districts in the American South. In the mid-twentieth century, following the desegregation Supreme Court cases described in Chapter Two, magnet programs emerged as "palatable integrators" across urban districts in the South, incentivizing White families to attend desegregated schools by offering enriching academic opportunities (Blank et al., 1996; Blank & Archbald, 1992; Frankenberg & Le, 2008; Frankenberg & Siegel-Hawley, 2010; Poppell & Hague, 2001;

Rossell, 2003; West, 1994). Currently, the majority of magnet programs still exist in urban school districts (Wang & Herman, 2017). Although magnet programs are not limited to the South nor to urban districts, Southern, urban school districts arguably still serve as the archetypes of magnet settings.

With this in mind, the setting for this research is Texas's major urban school districts. Texas is the most populous state in the South (United States Census Bureau, 2019), with five substantial metroplexes (Houston, Dallas-Fort Worth, Austin, San Antonio, and El Paso). The Texas Education Agency codes districts based on their enrollment and location, with 11 districts qualifying as "major urban" districts (Texas Education Agency, 2017). Collectively, these 11 served over 950,000 students in the 2018-2019 school year (Texas Education Agency, 2019). The demographic data of these districts are close to that of the state as a whole, although these districts serve higher proportions of Hispanic students, economically disadvantaged students, and English learners than Texas as a whole (Texas Education Agency, 2019). Table 3.1 displays demographic data for each district as well as for Texas as a whole.

This setting is appropriate for the research questions posed because it captures close to one million students within urban districts, the primary location for magnet programs, while limiting the sample to one state. Analyzing data from only one state enables consistency across schools and districts in terms of standardized tests, demographic data operationalization, and underlying policies and practices such as state and federal funding. In short, this purposive sampling balances a large sample size with consistency across variable operationalization, enhancing validity.

Table 3.1

Texas Major Urban Districts, 2018-2019 Demographics

District	Enrollment	Number of Schools	Number of Magnets	% African American	% Hispanic	% White	% Economically Disadvantaged	% English Learners
Arlington ISD	59,783	77	8	25.3	46.9	18.7	72.1	29.0
Austin ISD	79,787	130	17	6.9	55.3	29.8	53.0	28.2
Dallas ISD	155,030	232	27	21.6	69.8	5.8	85.9	45.6
El Paso ISD	57,178	92	13	3.3	83.5	9.5	73.9	31.9
Fort Worth ISD	84,332	143	27	21.7	63.4	11.3	84.1	34.2
Houston ISD	209,040	279	111	22.7	62.4	9.2	79.3	34.0
North East ISD	64,850	75	8	7.4	60.1	24.6	49.5	14.3
Northside ISD	105,797	120	9	6.7	67.8	18.6	47.8	9.9
San Antonio ISD	48,720	100	37	6.1	89.9	2.8	89.4	20.7
Socorro ISD	46,618	50	4	2.4	92.0	3.8	73.5	25.0
Ysleta ISD	41,036	59	10	1.4	94.7	3.1	78.3	28.7
Total of Sample Districts	952,171	1,357	271					
Average of Sample Districts	86,561	123	25	11.4	71.4	12.5	71.5	27.4
Total of Texas Public Schools	5,493,940	Over 9,000		12.6	52.8	27.0	60.2	20.3

Source: Texas Education Agency (2019).

Participants

Over 950,000 students were served by the districts included in this research's sample in 2018-2019, although school-level data will be analyzed (Texas Education Agency, 2019). In total, these districts consisted of 1,357 schools in the year analyzed (Texas Education Agency, 2019). Due to the Texas Education Agency's school-level data used in this study, individual students are not identifiable, in compliance with the Family Educational Rights and Privacy Act. FERPA policies significantly - and rightfully - limit access to individual student data. Thus, in this school-level analysis, I strive to capture overarching aspects of the learning opportunities of close to one million Texas students within a sample of 1,357 units - schools. This aggregation inherently lacks the specificity possible within a student-level study, but, given limited data accessibility, a school-level analysis provides a more granular perspective than other alternatives, such as using school districts or states as the unit analysis.

Dataset

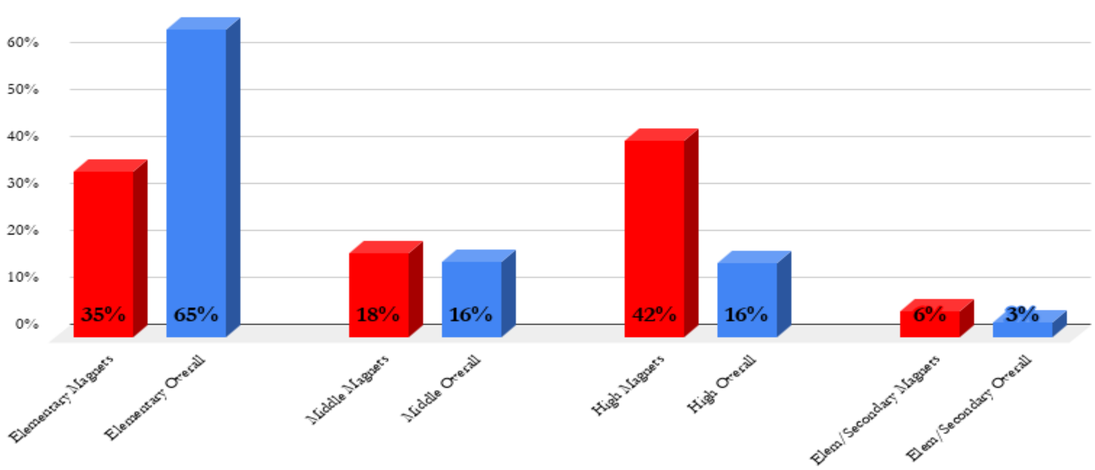
Description of Schools

The 1,357 schools within this study's sample vary significantly. Together, this dataset includes 886 elementary schools, 217 middle schools, 213 high schools, and 41 combination elementary/secondary campuses. I exclude 38 schools from my analysis due to their label as "alternate education" sites, meaning that they do not participate in state standardized testing due to special circumstances. An example is Houston ISD's Community Services School, which provides educational services to students who are long-term hospital inpatients or prison inmates. Further, 42 additional schools within the sample exclusively educate young learners between preschool and second grade. Since state standardized testing begins in third grade in Texas, those 42 campuses must be dropped from analysis, as they do not offer dependent variable data points (described below). 1,277 schools remain, after I account for these 80 campuses that do not participate in standardized testing.

This sample consists of 271 schools offering magnet programs, representing 20% of the full sample. In the section "Independent Variables," I describe my process for determining which campuses can be considered magnets in the context of this study. Among the magnets, 95 are at elementary schools, 48 at middle schools, 113 at high schools, and 15 at elementary/secondary combination schools. Magnets are disproportionately offered at high schools within this sample, as shown in Figure 3.1. Specifically, high schools make up 16% of the overall number of schools in the broader sample, but magnet high schools represent 42% of magnets studied. This contrasts with an underrepresentation of magnets in the elementary context, also visible in Figure 3.1. Although 65% of schools within the greater sample are elementary schools, a mere 35%

Figure 3.1

Percentages of Schools and Magnets within the Sample, by School Level



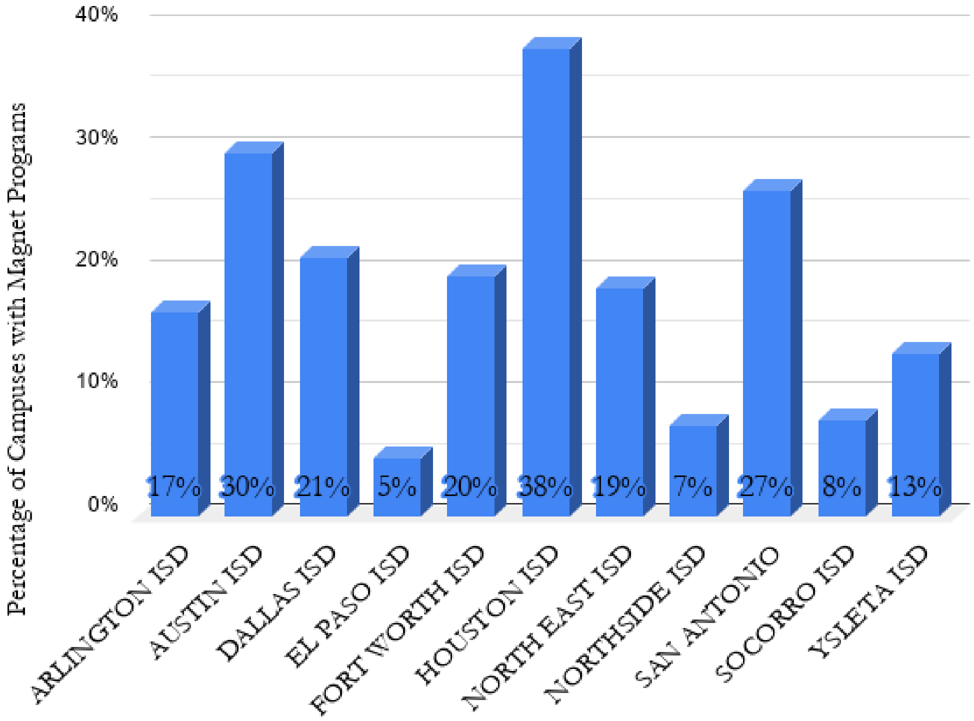
Source: Texas Education Agency (2019).

of the magnet programs analyzed are offered at elementary campuses. This discrepancy in magnet offerings across school levels may be due to a number of factors, such as stakeholder demand, and should be kept in mind when considering this work's findings.

The 11 major urban public school districts in Texas, which comprise my sample, each offer magnet programs. Unsurprisingly, the largest district - Houston ISD - offers the most, with 107 magnet schools. Houston ISD also has the highest percentage of magnets across the sample, as 38% of its schools offer magnet programming. On the opposite end of the spectrum, El Paso ISD has a mere four magnet programs, representing 5% of its campuses. Figure 3.2 displays the percentages of campuses with magnet programs at the 11 districts studied.

Figure 3.2

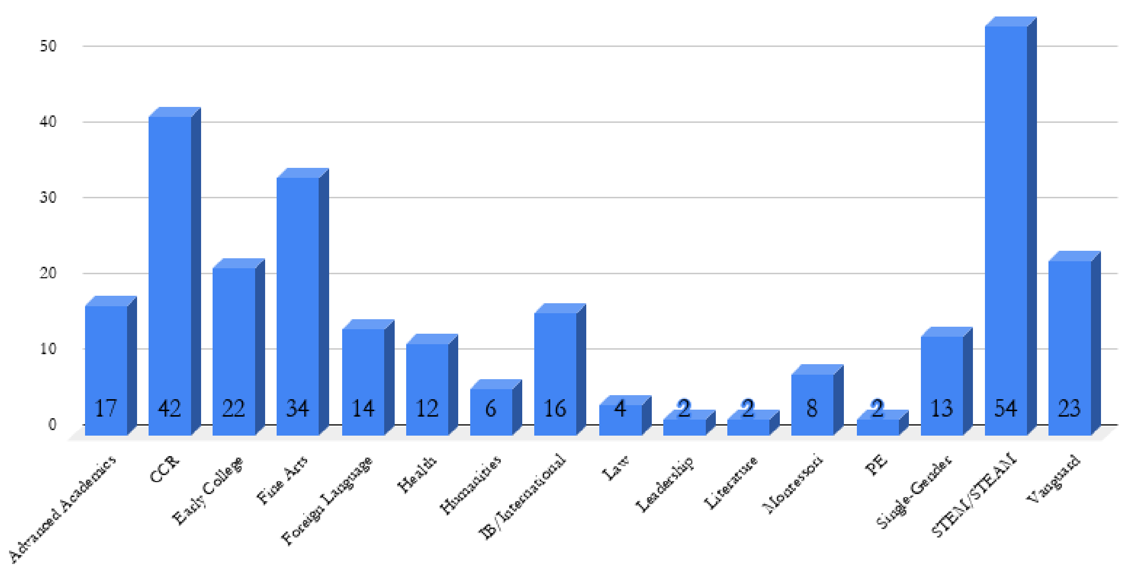
Percentages of Campuses Offering Magnet Programs, by District



Additionally, the 271 magnets in this sample offer a total of 16 pedagogical or curricular themes, as shown in Figure 3.3. As with magnet status, I articulate my process for determining magnet theme in the “Independent Variables” section below. These 16 themes are: advanced academics (rigorous instructional approaches such as problem-based learning), college and career readiness (in which students complete a technical certification or associate’s degree during high school), early college (in which students attend a local college and graduate high school with an associate’s degree), fine arts, foreign language, health, humanities, International Baccalaureate (IB) or international studies, law, leadership, literature, Montessori, physical education, single-gender, STEM/STEAM (science, technology, engineering, [arts,] and mathematics), and Vanguard (for students who have been identified as gifted and/or talented). Theme frequency ranges from 54 STEM/STEAM magnets to two offering leadership, literature, or PE magnets. Certainly, the dataset includes a broad, diverse range of schools.

Figure 3.3

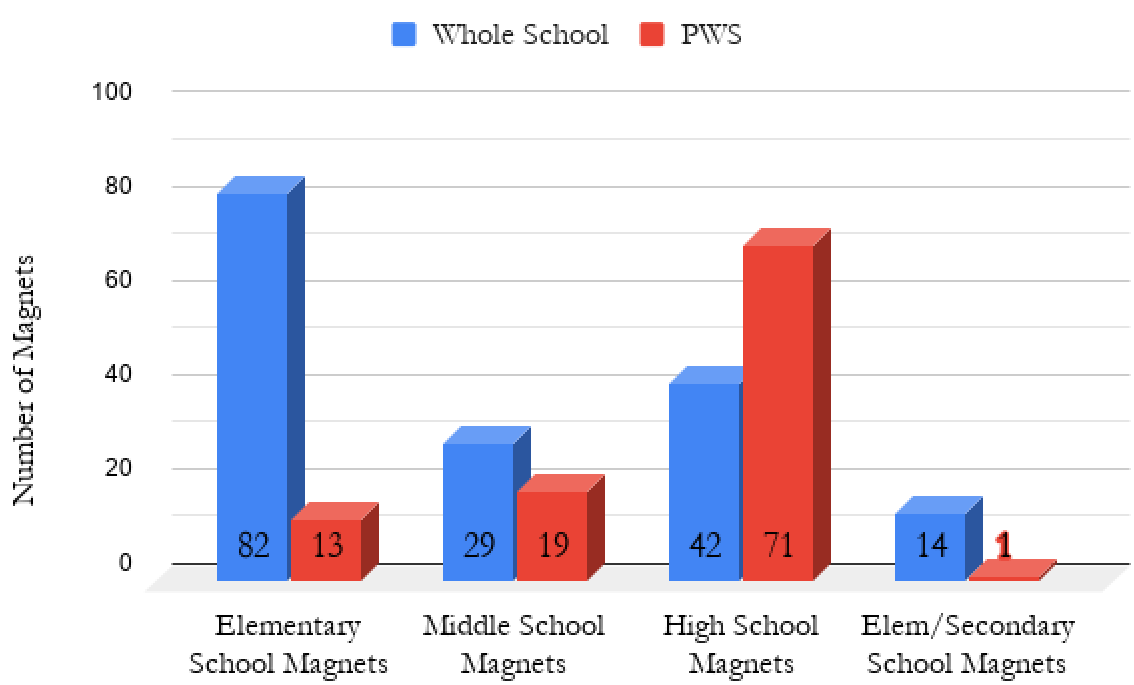
Magnet Theme Frequencies



Lastly, the magnet school subset of this sample includes both whole school and program-within-school (PWS) magnets. 167 of the 271 magnet programs within the Texas major urban school districts are whole school programs, including all students on campus in the magnet. The remaining 104 campuses offer PWS magnets, in which a subset of students participate in magnet programming. As shown in Figure 3.4, elementary magnets tend to use a whole school format, while PWS magnets are more prevalent among high schools in the sample.

Figure 3.4

Magnet Program Formats by School Level



Source of Demographic and Academic Data

In this study, I utilize data from the Texas Education Agency's publicly available warehouse of demographic, academic achievement, personnel, and financial data for each public school and district in the state. The Agency appropriately boasts that this dataset is "a treasure trove of information for researchers, parents and the public at large" (Texas Education Agency, n.d.c). Specifically, I download and analyze the Agency's TAPR reports (Texas Academic Performance Reports) for each of the 11 districts and 1,357 campuses within this purposive sample. Key demographic data I compile from these reports include each campus's enrollment and the percentages of students who are labeled as economically disadvantaged, English learner, receiving special education services, African American, Hispanic, and White. In terms of academic data, I tabulate the percentage of African American, Hispanic, and White students who pass key state standardized tests (fifth grade reading and math as well as English I, English II, and Algebra end-of-course exams). I record these percentages separately at each of the three passing performance levels - approaches grade level, meets grade level, and masters grade level. Additionally, I note each school's level (elementary, middle, high, or elementary/secondary) and what grades it offers, as well as identifying the alternate education campuses described above. These reports also provide me with each campus's average ELA and math class sizes, the principal's years of principalship experience, the teachers' average years of teaching experience, and the percentage of teachers with masters or doctoral degrees. I utilize 2018-2019 data for this research because the state standardized test, the State of Texas Assessments of Academic Readiness (STAAR), was not administered in the spring of the 2019-2020 school year due to the Covid-19 pandemic.

Sources of Magnet Status, Theme, and Format Data

The Texas Education Agency does not publish - nor appears to catalog, for that matter - schools' magnet status (whether or not a campus offers a magnet program) or program details. To garner this information, I access magnet program lists published on each of the 11 districts' school choice websites. While certain districts use different labels to describe magnet schools, including "specialty schools" and "academies," I consider all schools that can be found to fit the following definition as magnets for the purpose of this research: a public school "that offers a special curriculum capable of attracting substantial numbers of students of different racial backgrounds" (U.S. Department of Education, 2017, para. 2). As needed, I conduct follow-up searches of individual campus websites to clarify magnet status. Through these district and campus websites, I also identify each magnet school's theme and format (whole school or program-within-school). This case-by-case analysis, while time-intensive, enables me to create a resource not currently available through state sources - a dataset of the 1,357 campuses in Texas's 11 major urban districts, coded by magnet status, theme, and format with consistent variable operationalization across districts. The following section elaborates on how I operationalize and organize variables in this work.

Variables

Dependent Variables

Through each research question, I aim to capture at least one facet of the relationship between magnet status and within-school racial disparities in academic achievement. In this section, I discuss the operationalization of academic achievement and racial achievement gaps that I utilize in this work.

Academic Achievement. Texas's standardized testing program is entitled the State of Texas Assessments of Academic Readiness, or "STAAR." This set of criterion-referenced tests measures students' mastery of the Texas Essential Knowledge and Skills (TEKS), the state's academic standards. With few exceptions, third through eighth grade students enrolled in Texas public schools participate in STAAR testing each spring in key subjects (Texas Education Agency, n.d.a). Students test in math and reading each year in third through eighth grade, with additional tests in writing in fourth and seventh grade, science in fifth and eighth grade, and social studies in eighth grade. At the high school level, students must pass five end-of-course (EOC) exams within the STAAR umbrella - English I, English II, Algebra I, Biology, and U.S. History, which are requirements for graduation. Based on their scores, students receive a performance level label of "masters," "meets," "approaches," or "did not meet" grade level. These performance levels overlay one another, such that a student who "masters" their grade level also counts as having "met" and "approached" grade level. Performance levels, rather than raw score or percentage accuracy, serve as the primary reporting metric for STAAR performance. For all STAAR assessments, the "approaches" performance level constitutes a passing score. These data are published through the Texas Education Agency's TAPR reports, both for schools' and districts' complete student bodies and disaggregated by demographic labels.

While standardized test data are far from a perfect measure of students' educational experiences - a limitation I explore further in Chapter Five of this thesis - STAAR scores serve as the primary metric by which a number of stakeholders measure the efficacy of Texas public schools and districts. These stakeholders include policymakers, oversight agencies, district leaders, campus leaders, prospective and current families and students, and education researchers. Additionally, high school students must pass all five EOC STAAR assessments in order to graduate, further solidifying these assessments as the standard for academic proficiency and the gatekeepers to postsecondary education in the state.

In this research, I analyze math and reading STAAR scores at the fifth grade and high school (EOC) levels. I select math and reading because these are often regarded - whether implicitly or explicitly - as the priority academic subjects in K-12 education, with significant implications for students' promotion from grade level to grade level as well as entrance to postsecondary education. For this reason, although many magnet themes aim to enrich students' learning beyond these two subjects, I measure disparities in academic outcomes as measured by math and reading test scores. For opportunity gaps to narrow, students must receive opportunities to learn subjects with sufficient cultural capital to open doors in terms of college and career access. In order to serve students in the short and long term, magnet programs must balance any curricular enrichment (such as fine arts) with meaningful opportunities to learn high-yield academic content, such as math and reading.

Additionally, I select fifth grade and EOC exams because these grade levels typically serve as the culmination of a student's tenure at an elementary or high school campus, thus maximizing the amount of time a student was exposed to that school's educational opportunities. Fifth grade math and reading STAAR scores also serve as promotional standards, meaning that public school students cannot transition to sixth grade without passing both assessments (absent an overriding determination from a grade placement committee). Similarly, this holds true for EOC exams and high school graduation. It should be noted that I average English I and English II scores for composite "English EOC" variables to facilitate data analysis and interpretation. Recall that I analyze 2018-2019 test scores in this work, as 2020 spring STAAR assessments were cancelled due to the Covid-19 pandemic.

Student Race. The other element necessary to quantify the racial achievement gap is student race. In Texas, students' parents or guardians select race/ethnicity labels when enrolling their children in public schools. The labels currently available are African American, Hispanic, White, American Indian, Asian, Pacific Islander, and Two or More Races. These data are warehoused in the Texas Education Agency's student data management system, the Public Education Information Management System (PEIMS). These data populate into the TAPR reports published by the TEA. As a reminder, these data are accessible at the school, district, and state level, so demographically disaggregated STAAR reports communicate average performance among students labeled with a given racial identifier.

The Gap. Understanding the operationalization of academic achievement and student race in the context of this thesis, these components can now be leveraged to quantify the racial achievement gaps measured in this work. The unofficial industry standard of racial achievement gap reporting is that which the U.S. Department of Education uses: separately measuring White/African American and White/Hispanic disparities in academic performance. In keeping with this practice, I measure White/African American and White/Hispanic STAAR score gaps at the masters, meets, and approaches performance levels on the key STAAR assessments detailed above.

Although gaps may exist between other demographic groups, less than eight percent of Texas public school students in 2018-2019 identified as a race other than White, African American, or Hispanic (Texas Education Agency, 2019), such that an impractically small number of students outside of these groups may be present at a given campus in this sample - such as Pacific Islanders in the fifth grade at a certain school. In fact, FERPA prohibits state education agencies from publishing data from such small student groups that individuals could potentially be identified, so STAAR performance from small student groups is excluded from TAPR report, marked with an asterisk.

For the purpose of this thesis, I quantity the racial achievement gaps within the sample as 1) the difference between the percentage of White students who master/meet/approach grade level on a given STAAR test and the percentage of their African American peers who do so and 2) the difference between the percentage of White students who master/meet/approach grade level on a given STAAR test and the percentage of their Hispanic peers who do so. Again, I quantify this gap on fifth grade reading, fifth grade math, English EOC, and Algebra EOC assessments. These calculations result in continuous variables for each disparity with a potential range of -100 (0% of White students passing at a given level while 100% of African American or Hispanic students pass at that level) to 100 (0% of African American or Hispanic students passing at a given level while 100% of White students pass at that level).

The concept of measuring racial achievement gaps inherently presents limitations, as no quantitative metric can thoroughly capture the disparities in learning opportunities and students' lived experiences. While I discuss those limitations in Chapter Five, I aim to balance specificity, thoroughness, and comprehensibility by education practitioners and researchers in my operationalization of racial achievement gaps for the purpose of this thesis. Thus, I employ the unofficial industry standard for reporting academic achievement among Texas public schools (STAAR scores), family-generated racial identifiers, and the federal government's precedent for achievement gap quantification.

Independent Variables

In research question 1, the primary independent variable of interest is magnet status - whether or not a campus offers a magnet program. The independent variable of question 2 shifts to magnet theme. I describe the operationalization for both below.

Magnet Status. To answer research question 1, I determine which of the 1,357 schools within Texas's 11 major urban public school districts offer a magnet program. As described above, I garner this information by visiting each district's school choice webpage, following up with individual campus websites. I label each school "that offers a special curriculum capable of attracting substantial numbers of students of different racial backgrounds" (2017, para. 2) as a magnet for the purposes of my analyses.

To recognize and account for the significant difference between whole school and program-within-school (PWS) magnets, I code these two categories separately. To reiterate, in whole school programs, every student on campus participates in the magnet, while PWS magnets limit programming to a subset of the student body (Goldring, 2009). I use the previously referenced district and campus websites to determine which magnets fall into each category then code the whole school magnets in the sample as "1" and the PWS magnets as "2." Additionally, I code schools without magnet programs ("nonmagnets") as "0." Stata 17 recognizes this 0/nonmagnet label as the reference category, such that the regression coefficients of the whole school magnet and PWS magnet variables will represent a value added for magnet programming, specific to the format.

Magnet Theme. To answer research question 2, I replace the magnet status variables with magnet theme variables in order to distill the nuances between distinct curricular and pedagogical programs. As described previously, I base these determinations on information on district and school websites. Each theme within this categorical variable receives a numerical code (1-16) to simplify data analysis.

To maintain subsamples of significant sizes and streamline interpretation of results, I cluster related categories together. For example, I combine STEM (science, technology, engineering, and math) magnets with STEAM (science, technology, engineering, arts, and math) magnets. I group foreign language magnets of various languages together as well, as I do for International Baccalaureate (IB) magnets and international studies magnets without a formal IB designation. Additionally, I recognize single-gender magnets that serve male or female students as one group. Approximately five campuses offer two magnets. For these unique cases, I code the school based upon the magnet theme that appears to have a larger presence in the school based upon program enrollment and/or its representation on the school website. The 16 themes present within the sample are described above in the "Description of the Schools" section.

Control Variables

To explore the relationship between magnet programs and their schools' racial disparities in standardized test scores, I implement nine control variables. These are: school district, campus enrollment, the percentage of students who qualify for free or reduced-priced lunch, the percentage of students who are English learners, the percentage of students who receive special education services, average class size, principal's years of experience, teachers' average years of experience, and the percentage of teachers with advanced (master's or doctoral) degrees. As a reminder, I aggregate these data from the 2018-2019 TAPR reports from the Texas Education Agency. I explain the justification and operationalization of each of the control variables below.

School District. In Texas, school districts are empowered with a great deal of autonomy due to the state's decentralized public education model. Each district receives some degree of independence in central issues such as budget appropriations, student assignment policies, magnet program funding, and magnet application practices. These factors, in addition to countless others, can directly and indirectly influence opportunities to learn. For example, in the 2018-2019 school year, San Antonio ISD spent $15,168 per student (Texas Education Agency, 2020a), while Socorro ISD spent $11,576 (Texas Education Agency, 2020b). This disparity may impact the educational inputs the average San Antonio ISD student receives compared to one in Socorro ISD. To address these differences, I create dummy variables to capture simple fixed effects for each district.

Campus Enrollment. Within this sample, school enrollments range from one student to 5,680 students per campus. To ensure that racial achievement gaps at each campus are considered proportionately, I control for campus enrollment throughout my analyses. Controlling for campus enrollment enhances the proportionality and the mathematical influence of individual campuses' data points, as a given elementary school may serve 100 students or 1,000 students.

Percentage of Students Who Qualify for Free or Reduced-Price Lunch.

Students living in poverty or economic instability disproportionately experience a number of disadvantages, including food insecurity and inadequate nutrition, increased physical and mental health conditions related to and exacerbated by less access to healthcare, chronic stress, exposure to environmental dangers, substandard housing, greater likelihood of homelessness, and increased exposure to violence (American Psychological Association, 2016; Aratani, 2009; Buckner et al., 2004; Jensen, 2009). These issues are not left at the schoolhouse gate when children arrive to campus but rather impact their learning in significant, measurable ways. For example, researchers have identified correlations between poverty in children and under-resourced schools, delays in cognitive development, below-average memory and concentration, externalizing problem behaviors, lower test scores, higher dropout rates, and increased discipline rates (American Psychological Association, 2016; Elliot, 2013; Howard, 2011; Jensen, 2009; Shonkoff & Garner, 2012). Certainly, poverty impacts children's lives, including their opportunities to learn.

Practitioners and researchers have historically operationalized poverty through a common yet imperfect measure - the percentage of students who qualify for free or reduced-priced lunch at school. At the beginning of each school year, school districts assign paperwork to enrolled families, including a form asking for information regarding the income of adults in the home. Eligibility thresholds are determined annually by the federal government's Department of Agriculture (United States Department of Agriculture, 2018). During the 2018-2019 school year, children in families with incomes within 130% of the federal quantification of "poverty" qualified for free lunch, and those within 185% of the poverty cutoff qualified for reduced-price lunch (United States

Department of Agriculture, 2018). This measure of students' socioeconomic experiences is far from perfect, with issues including incomplete and inaccurate paperwork, as well as the aggregation of reduced-price lunch rates and free lunch rates into a single statistic (Snyder & Musu-Gillette, 2015). However, as this proxy is the norm in the field and represents the accessible data regarding student poverty, I use free and reduced-price lunch percentage in this research to operationalize the number of students in each school studied who must navigate the complicating life factors described above.

Free and reduced-price lunch percentage must be controlled in this research due to its significant variance across campuses in the sample, as students living in poverty are not distributed equally or equitably across schools. Within the 11 districts studied, schools have free or reduced-price lunch rates as low as 2.3% at Clayton Elementary School in Austin ISD and as high as 100% at over three dozen schools in the sample (Texas Education Agency, 2019). As of 2019, Houston ISD alone contained 19 schools where 100% of students met the threshold for this service. This heterogeneity justifies controlling this variable in this study.

Percentage of Students Who Are English Learners. Select students across the state of Texas, in addition to learning the Texas Essential Knowledge and Skills (TEKS standards), are also developing multilingualism as they learn English. Across the state, all English learners and their teachers must balance these dual priorities during the school day. This requires that educators foster specific pedagogical skills and navigate competing time demands (Austin-Archil, 2019; Avila, 2015; Gandara & Hopkins, 2010; Helfrich & Bosh, 2011; Robertson, n.d.; Rubin, 2016). This proves highly relevant in Texas, where, in 2017, 18% of students were labeled English language learners (National

Center for Education Statistics, 2017). English learners experience distinct needs, and the many educators across Texas who serve them must ensure instructional excellence and devote time to support them as they develop multilingualism.

As with free or reduced-price lunch qualification, students are classified as English learners based upon enrollment paperwork. Families are asked to report the language typically spoken in the home, the language the student speaks most of the time, the student's place of birth, and the number of years of school in and out of the United States. Based on this information, schools identify students who may benefit from explicit instruction in English listening, speaking, reading, and writing. The English learner percentage used in Texas, including in the TAPR reports, is operationalized as the "percentage of students whose primary language is other than English and who are in the process of acquiring English. The terms 'English learner' and 'Limited English Proficient' (LEP) are used interchangeably" (Texas Education Agency, n.d.b). I access the percentage of English learners in each district and campus from the TAPR reports described throughout this chapter.

Additionally, as with poverty, English learner status varies significantly across the sample, justifying the methodological choice to control this variable. Some schools within the sample serve no documented English learners, while others exclusively enroll English learners. For example, several major urban districts offer schools targeted to meet the needs of students who have recently immigrated to the United States. These include the Newcomer School in Arlington ISD (98.1% English learners), the International Newcomer Academy in Fort Worth ISD (98.4%), and Las Americas in Houston ISD (100%) (Texas Education Agency, 2019). Again, this significant heterogeneity justifies the choice to control the variable.

Percentage of Students Receiving Special Education Services. Students receiving special education services may face above-average barriers to success on standardized tests, depending upon their disability. Compared to nondisabled peers, students with disabilities may struggle with comprehension, processing, attention, memory, and other skills reflected on standardized tests, which can contribute to the achievement gap between students who do receive special education services and those who do not (Eckes & Swando, 2009; Schulte & Stevens, 2015; Thurlow et al., 2016; Wei et al., 2012). It should be noted that, as with English learners, students with disabilities can and do learn, grow, and succeed with teachers who tailor their instruction to build on students' strengths using pedagogical techniques aligned to students' learning styles and needs (Hurwitz et al., 2019; Mintrop & Zane, 2017).

The percentage of students receiving special education services merits inclusion as a control variable for the same reasons as free and reduced-price lunch and English learner percentages. Namely, this factor can impact students' learning opportunities, and it is characterized by significant heterogeneity within the sample. Special education percentage tends to fluctuate between 5% and 15% in Texas's major urban districts, but the range of this variable represents an extremely broad spectrum from 0% to 99%. For example, none of the 407 students at Arlington Collegiate High School in Arlington ISD receive special education services (Texas Education Agency, 2019), a potential equity concern for this early college magnet. On the opposite end of the spectrum, 99% of the 141 students enrolled at the Rosedale School in Austin ISD, a K-12 campus for students with significant disabilities, do (Texas Education Agency, 2019). The presence of such extreme outliers among this influential factor justifies controlling special education percentage, similar to the rationale for controlling free and reduced-price lunch and English learner percentages.

Average Class Size. An additional variable that has the potential to influence learning opportunities and can vary significantly across campuses is class sizes. Practitioners often anecdotally cite class size as a substantial factor in teaching and learning, but the empirical research suggests that this factor's influence is limited. In his landmark meta-analysis of conditions influencing student achievement, Hattie (2009) found that class size had a present but low impact on achievement, with an average effect size of $d = 0.21$. This finding aggregated the results of 96 studies and three meta-analyses, which I do not report individually here for the sake of brevity. While the effect size is low, it is not insignificant and thus could potentially influence the relationship between magnet programming and within-school racial achievement gaps, justifying its inclusion in these models.

Class size, as with most of the variables incorporated in my research, is reported in TEA's TAPR reports. Specifically, each campus's report lists kindergarten through fifth grade average class size and, at the secondary level, average math class size and average English/language arts class size. These values are derived from district records outline teachers' course responsibilities and offer a quantitative indicator of this aspect of students' learning environment.

Principal's Years of Experience. Principals ideally serve as the central instructional leader for their schools, shaping instructional practices and learning environments. Among the greater field of education research, a significant portion of the extant literature addresses school leadership. In fact, Hattie's (2009) meta-analysis included a staggering 491 studies and 11 meta-analyses exploring the influence of principals or other school leaders on student achievement. He calculated an overall impact of $d = 0.36$, narrowly missing his 0.40 cutoff for a medium effect size and inclusion in his "zone of desired effects." This suggests that campus leadership is significantly associated with student outcomes but certainly cannot serve as a cure-all that could compensate for insufficiencies in other key variables. Again, I do not detail the studies Hattie (2009) analyzed and lean upon his aggregation of the extant literature for the sake of brevity.

As outlined in Chapter Two of this thesis, researchers have explored the influence of magnet program leadership specifically, although this research avenue has been less prevalent in recent years (Bauch & Goldring, 1996; Blank et al., 1996; Christenson et al., 2003; Chubb & Moe, 1990; Clewell & Joy, 1990 Dentler, 1990; Frankenberg et al., 2008; Hausman, 2000; Hausman et al., 1997; Hausman & Goldring, 2001; Smrekar & Goldring, 1999; Walton & Ford, 2014; Wehlage & Smith, 1992). Altogether, it appears that principal leadership may impact students' learning opportunities and outcomes in the magnet context.

Although principal leadership is a complex, multi-faceted concept, the available data offer limited metrics. The TAPR reports include each campus principal's years of experience as a principal and years of experience as a principal within their school district. These statistics are reported for each school's assistant principal(s) as well. For the purpose of this study, I incorporate the statistic capturing principals' total years of principalship experience based upon the logic that a principal with more experience in the role may have stronger instructional leadership skills that could benefit the learning opportunities their campus provides. It could be argued as well that more experienced principals also shape learning environments through their expertise related to other factors, such as student discipline. In this sample, this variable ranges from zero years of principalship experience to 24. The potential importance as well as this heterogeneity suggest that principal experience is a worthwhile control variable to include in these regression models.

Teachers' Average Years of Experience. Similarly, and perhaps even more importantly, teaching experience may influence students' learning opportunities and outcomes. A rigorous literature review of 30 recent studies (Kini & Podolsky, 2016) found that more years of teaching experience correlated with greater student achievement gains as well as additional academic benefits, such as improved student attendance. Although not every experienced teacher offers better learning opportunities than every novice teacher, it stands to reason that, over the course of their careers, teachers develop increasing instructional expertise that can translate to enhanced student learning.

As with principal experience, TEA's TAPR reports include values related to teaching experience, with greater detail than the reports' principalship statistics. Each TAPR report quantifies the number and percentage of teachers on the campus with zero, one to five, six to 10, 11 to 20, and more than 20 years of experience. Additionally, it includes the average years of experience of the school's teachers as well as their average years of experience within their school district. I incorporate the average years of teaching experience variable in my model, comparable to the average years of principalship experience variable. This statistic offers an at-a-glance value capturing the degree to which a given school's faculty has developed experience in their field. As with the principal experience data, this study's dataset includes a wide spectrum in terms of teaching experience, ranging from zero to 36 years. Based upon the potential importance of teaching experience on student learning and achievement, as well as the variability of this potential resource across campuses, I include this learning input variable in my regression models.

Percentage of Teachers with Advanced Degrees. The final control variable I incorporate in my models is the percentage of teachers with advanced (master's or doctoral) degrees. The extant literature indicates that advanced degree programs have the potential to benefit teachers and their students, although these findings are inconsistent and suggest that this factor can vary significantly by context (Barnett, 2020; Campbell & Lopez, 2008; Chang et al., 2020; Ladd & Sorensen, 2015). Additionally, Ladd and Sorensen (2015) found that teachers with master's degrees were associated with improved student attendance compared to teachers without advanced degrees, suggesting that advanced degrees may indirectly influence student learning. Again, these findings are contextually bound, and it appears that "the devil is in the details" (Campbell & Lopez, 2008, p. 33) in terms of the relationship between teachers' advanced degrees and student outcomes.

As with the other three control variables related to academic inputs – class size, principalship experience, and teacher experience – TEA publishes data regarding teachers' advanced degrees in the TAPR reports. The count and percentage of teachers whose highest degree is a bachelors, a master's, or a doctorate is listed in these reports. Additionally, the reports include a count and percentage for teachers without any degree, but this value is always zero due to the state's requirements for teacher licensure. I aggregate master's and doctoral percentages into a single "advanced degrees" statistic, which ranges from zero to 100 percent in this sample. As with many of the control variables I propose for this study, the potential impact of this factor as well as its heterogeneity within this sample justify its inclusion in my modeling.

Potential Control Variable Not Included. In addition, I wish to note a set of control variables I considered but have chosen not to include within this thesis – schools' racial percentages. I weighed the advantages and disadvantages of controlling for the percentage of each school's student population that has been labeled as African American and the percentage labeled as Hispanic. Students' racial labels, however, have already been captured in these regression models through the operationalization of the dependent variable. Incorporating this element a second time could potentially introduce a risk of simultaneity. Additionally, I did not want to decrease the weight of the data from students with small percentages of students of color. In my opinion as an educator and a researcher, the disparities in learning opportunities and academic achievement among schools where students of color are in the minority are just as important to measure and address at those where they are in the majority. Certainly, future researchers may wish to navigate this decision differently, an implication I address in Chapter Five.

Instrumentation and Procedure

Instrumentation and Dataset Aggregation

To address the research questions outlined above, I conduct secondary analyses of quantitative data from the Texas Education Agency. As described above, I access demographic and STAAR standardized test scores published in TEA's publicly available

Texas Academic Performance Reports (TAPRs). I compile the data from each of the 1,357 schools' TAPR reports in a GoogleSheet spreadsheet to aggregate all demographic and academic data points within a single dataset. Given the lack of personal identifiable information in this school-level, publicly available data, my security measures are limited to password-protecting my GoogleDrive.

Then, I conduct statistical analyses using Stata 17 (StataCorp, 2019). Specifically, I run multiple regression models to answer each research question. I describe my procedures for each analysis below.

The Case for Multiple Regression

In both of my research questions, I explore the relationship between magnet programs and within-school racial achievement gaps. I include nine control variables within both models to distill the effects of magnet programming on these gaps. Only one statistical test can measure the effect of a categorical independent variable on a continuous dependent variable while controlling for additional variables (both continuous and categorical): multiple regression (Bailey, 2015; Cohen et al., 2011; Lewis-Beck & Lewis-Beck, 2016). Multiple regression "enables us to predict and weight the relationship between two or more *explanatory* … variables and an *explained* … variable" (Cohen et al., 2011, p. 663). Compared to other analyses of difference such as t-tests and ANOVA, multiple regression offers the distinct advantage of accounting for control variables to distill the particular relationship of a given independent variable and the dependent variable. As I endeavor to evaluate the value added of magnet programming on students' opportunities to learn, multiple regression enables me to control for contextual variables to focus my analysis on the effects of magnet status on within-school racial achievement gaps.

Regression Assumptions

Multiple regression relies upon a number of assumptions regarding features of the variable, sample-level conditions, variable relationships, and residuals, which Cohen et al. (2011) outline. The first two assumptions relate to variable characteristics. Namely, the dependent variable must be a continuous variable, and there must be at least two independent variables. These assumptions are inherently met due to the variable operationalization as described above. Specifically, the difference in standardized test scores is a continuous, ratio variable, and I utilize several independent and/or control variables (magnet status, enrollment, district, magnet format, student demographic percentages, and magnet themes). No tests are needed to verify these assumptions.

Next, regression requires several sample-level conditions (Cohen et al., 2011; Lewis-Beck & Lewis-Beck, 2016). Multiple regression is quite sensitive to outliers, so an additional assumption is a lack of significant outliers, which could distort the model. To ensure I did not input any data errors while aggregating the dataset, I will run descriptive statistics on each of the variables to look for inappropriate values, including examining the minimums and maximums for each continuous variable. For example, all achievement gaps should fall between -100 and 100, as I calculate them by subtracting a number between 0 and 100 from another number between 0 and 100. Samples must also be free of autocorrelation in order to conduct multiple regression. This means that there should be no perfect correlations between individual cases, which I will check for using the runs test within Stata. Additionally, the mathematical bases for regression require that the sample contain more observations than parameters within the model. This is not a concern here, as my sample consists of 1,357 observations. I will any violations of the assumptions described in this paragraph within Chapter Four.

Regression also relies upon two key assumptions regarding the relationship between variables within the model (Cohen et al., 2011). First, because multiple regression extends the logic of traditional linear regression, there must be an approximately linear relationship between the dependent variable and any continuous independent variables, as well as between the dependent variable and the continuous independent variables together. I will examine scatterplots for each dependent-independent variable relationship to evaluate whether any variables must be transformed to approximate linearity. Another central assumption of multiple regression is an absence of perfect multicollinearity, meaning that none of the independent variables included in the model are perfectly correlated with one another. To check for multicollinearity, I will use Stata to run VIF (variance inflation factor) tests. However, finding a degree of multicollinearity will not invalidate this model, only increase its standard error (Achen, 1982). Again, I will report any violations of these assumptions in Chapter Four.

Finally, Cohen et al. (2011) describe assumptions regarding the residuals of multiple regression models. Specifically, these residuals must be normally distributed with a mean of 0, have equal and constant variance, and not be correlated with any of the independent variables in the model. Within Stata, I will generate a residual variable for each model under analysis to enable statistical analysis of these error terms. I will run normality tests on the residuals to evaluate the skewness and kurtosis of their distribution. Simple descriptive statistics will allow me to determine if the mean of each residual variable is 0. To check for homoscedasticity - the presence of equal, constant variance among the residuals - I conduct the Breaush-Pagan and Cook-Weisberg tests. If I uncover heteroscedasticity, I will address this issue using Stata's robust standard errors option. Lastly, to ensure residuals are not correlated with any independent variables, I will run correlation tests in Stata between the residuals and the independent variables. Through these processes, I will ensure the residuals fulfill multiple regression's assumptions, and I will report any violations of these assumptions in Chapter Four.

Regression Models

I can explore both of my research questions using a distinct set of regression models. Within-school racial achievement gaps remain the dependent variable in both models, and the control variables remain the same as well. However, the independent variable varies by question. I quantify White/African American achievement gaps by subtracting the percentage of African American students who score at a particular performance level on a given STAAR test from the percentage of their White peers who do so. Similarly, I calculate White/Hispanic gaps by subtracting the percentage of Hispanic students who score at a particular performance level on a given STAAR test from the percentage of their White peers who do so. I will conduct analyses for both the White/Hispanic and the White/African American gap at each of the three passing performance levels (masters, meets, and approaches grade level) and each of the four STAAR tests studied (fifth grade math, fifth grade reading, English I and II, and Algebra I). In total, 24 gaps will be explored through this methodology when considering the combinations of racial gap (two), performance levels (three), and tests (four). This granular approach enables the potential discovery of results that could otherwise be hidden with a broader, less specific analysis. This increased potential of significant findings supports my overall goal of identifying actionable opportunities for school districts to better serve their students. For example, if the only significant findings from my second research question is that a lower achievement gap exists between White and African American students passing the fifth grade math test and between White and Hispanic students meeting expectations on the fifth grade reading test, this would provide guidance that districts should target their magnet programming at the elementary school level. Granularity and increased opportunities to uncover significant findings, in this case, align with the true mission of this thesis and, frankly, my doctoral journey - uncovering avenues to minimize the racial opportunity gap in Texas's major urban school districts.

Unlike the dependent and control variables, the independent variable varies across the two questions. For the first question ("What is the nature of the relationship between magnet status and within-school racial achievement gaps, controlling for contextual variables known to influence academic outcomes?"), magnet status serves as the independent variable. The formula for this model is:

$$Y_{i\ racialgap} = \beta_0 + \beta_1 X_{PWSmagnet} + \beta_2 X_{wholemagnet} + \beta_3 X_{enrollment} + \beta_{4\text{-}13} X_{districtdummy} + \beta_{14} X_{FRPLpercentage} + \beta_{15} X_{ELpercentage} + \beta_{16} X_{SpEdpercentage} + \beta_{17} X_{classsize} + \beta_{18} X_{advanceddegrees} + \beta_{19} X_{principalexperience} + \beta_{20} X_{teacherexperience} + u_i$$

I adapt the model to address my second research question ("Does this relationship vary across magnet program themes?"). In this question, magnet theme replaces magnet status as the independent variable. Magnet theme is a categorical variable and thus, for the purpose of regression calculations, I must express it as a series of dummy variables (Bailey, 2015; Lewis-Beck & Lewis-Beck, 2016). The reference category for magnet theme is the absence of a magnet program, so regression coefficients will represent a value added of magnet programming. This formula is as follows:

$$Y_{i\ racialgap} = \beta_0 + \beta_1 X_{PWSmagnet} + \beta_2 X_{wholemagnet} + \beta_3 X_{enrollment} + \beta_{4\text{-}13} X_{districtdummy} + \beta_{14} X_{FRPLpercentage} + \beta_{15} X_{ELpercentage}$$

$$+ \beta_{16} X_{SpEdpercentage} + \beta_{17} X_{classsize} + \beta_{18} X_{advanceddegrees} + \beta_{19} X_{principalexperience} + \beta_{20} X_{teacherexperience} + \beta_{21\text{-}37} X_{magnetthemedummy} + u_i$$

Regression Reporting

In Chapter Four, I will report the findings of my multiple regression analyses. I will report both the goodness of fit and the coefficient results of each regression model in table and written form. I will share any regression violations that arise as well. Below, I describe this reporting.

To evaluate how well the regression models explain the variance within the dependent variable, I will report the *R* squared value for each model. This statistic "tells us how much variance in the dependent variable is explained by the independent variable in the calculation" (Cohen, 2011, p. 662). Specifically, the value represents the percentage of variance the model explains (Cohen, 2011; Lewis-Beck & Lewis-Beck, 2016). For example, an *R* squared of .387 would signify that the model under analysis explains 38.7% of the variance in within-school racial achievement gaps for that particular model. Mujis (2004) recommends that an *R* squared value between 0 and 0.1 be considered a poor fit, 0.11-0.3 a modest fit, 0.31-0.5 a moderate fit, and 0.51-1 a strong fit. The *R* squared statistic offers a well-controlled, readily interpretable representation of the goodness of fit of each of my multiple regression models.

Then, I will answer the research questions through reporting the coefficients of relevant Beta values within the regression model. Stata will calculate coefficients for each independent and control variable within the models, which quantify the impact each variable has on the dependent variable, all other independent and control variables held equal (Cohen, 2011; Lewis-Beck & Lewis-Beck, 2016). It should be noted that coefficients will be calculated in decimal form and require multiplication by 100 to express the change in percentage points of the relevant achievement gap. For example, in the context of an analysis of the White/Hispanic gap in eighth grade math masters percentage, a coefficient of 0.153 for the early college magnet variable would suggest the following interpretation: All else held equal, an early college magnet program predicted an increase in the gap of White and Hispanic students scoring masters on the fifth grade math STAAR test by 15.3 percentage points relative to nonmagnet campuses. One can see the benefit of using nonmagnets as the reference category for the magnet theme dummy variables in this example. To communicate the weight of each coefficient, I will also report the *t* statistic and statistical significance (*p* value) of each coefficient. Greater *t* statistics and lower *p* values represent stronger statistical significance. I regard a *p* value of less than 0.05 to be representative of statistically significance, aligned with the traditional threshold within education research. As with *R* squared, I will report these coefficients and their related statistics through both tables and sentences to offer the reader multiple avenues for reading and understanding meaningful results.

Conclusion

In this chapter, I have outlined my path for answering the two research questions I use to direct my exploration of the effects of magnet programs and their themes on within-school racial achievement gaps. First, I reiterated those questions, outlined the general research design, and shared the hypotheses for both research questions. I then described the sample for this work, including its setting, participants, and dataset. Following this, I detailed the variables incorporated in the models. Finally, I justified the use of multiple regression to answer my research questions, described my model specifications, and explained the process of conducting these tests and reporting their results. In Chapter Four, I share my findings as they pertain to the research questions of this thesis.

Chapter IV

Results

In this chapter, I present the results of the statistical analyses outlined in Chapter Three. I leveraged multiple regression to address both research questions in an effort to explore the relationship between magnet status or theme and within-school racial achievement gaps, controlling for relevant variables. I found numerous models as well as several magnet themes to be closely related to within-school racial achievement gaps at the $p < 0.05$ level. Before reporting these findings, I contextualize my analyses through relevant descriptive statistics. I then present the results of the regression models for each research question and conclude with a summary of the significant findings.

Descriptive Statistics

To provide important context to the subsequent regression results, in this section, I offer descriptive statistics regarding the models' variables. This information describes relevant characteristics of the schools incorporated within the sample as well as their racial achievement gaps. Here, I briefly summarize key features of the variables related to student demographics, educational inputs, and within-school racial achievement gaps at the 1,357 campuses under analysis within this thesis.

Demographics varied significantly among the schools in Texas's major urban school districts, as displayed in Table 4.1. Within this sample, the average campus served 702 students, of whom 77% had been identified as economically disadvantaged, 31% were English learners, and 11% received special education services. A broad range existed within each variable, reflective of the heterogeneity of schools offered across Texas's five metroplexes. The raw count of students enrolled at a given campus spanned from 1 to 5,680 students, and prevalence of each of the three student labels (economic disadvantage, English learner, and special education status) ranged from 0% to 100%. (It should be noted that I present percentages in decimal format in this and all further tables.)

Table 4.1

Descriptive Statistics of Control Variables

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Student Demographic Variables					
School enrollment	1,357	701.67	543.86	1	5,680
Economic disadvantage percentage	1,357	0.77	0.24	0	1
English learner percentage	1,357	0.31	0.22	0	1
Special education percentage	1,298	0.11	0.09	0	1
Educational Input Variables					
Average class size (5th grade)	825	20.08	5.61	1	49
Average class size (secondary ELA)	466	15.92	5.50	1	28
Average class size (secondary math)	473	17.20	6.03	1	32
Percentage of teachers with advanced degrees	1,330	0.27	0.12	0	1
Principal’s years of experience	1,330	6.02	4.80	0	24
Teachers’ average years of experience	1,330	11.00	2.99	0	36

The educational input variables incorporated in this thesis offered a great deal of heterogeneity as well, also shown in Table 4.1. The average fifth grade class size in this sample was 20 students, the average secondary English Language Arts class size was 16, and the average secondary math class size was 17. On the average, 27% of a campus's teachers had a masters or doctoral degree, and the mean for a faculty's average teaching experience was 11 years. Within this sample, the average principal had six years of principalship experience. Again, the ranges for each of these variables were substantial. For example, fifth grade, secondary ELA, and secondary math classes in the sample enrolled as many as 49, 28, and 32 students, respectively. The reported fifth grade class of 49 students at Locke Hill Elementary School in Northside ISD (TEA, 2019) was likely an outlier due to a special circumstance. For example, long-term substitute teachers are not incorporated in this TAPR statistic, affecting the denominator in these "students per teacher" calculations. Each class size variable had a minimum of one, a class size which occurred at disciplinary alternative education placements (DAEPs), juvenile justice alternative education placements (JJAEPs), and schools serving homebound students. In terms of faculty and administration, the range for teachers' advanced degrees was 0% to 100%, the range for principal experience was zero to 24 years, and the range for average teacher experience was zero to 36 years. A statistic of zero years of experience indicates that the 2018-2019 was the principal's first year in the role or that the average teacher was in their first year teaching, such that all teachers were new to the field. Interestingly, four schools in the sample had an average of zero years of teaching experience, each representing a special circumstance such as Houston ISD's JJAEP and Austin ISD's preschool serviced by United Way (TEA, 2019). It should be noted that many of the outliers do not serve a sufficient number of students participating in STAAR testing, so they are dropped from analysis in this thesis. Considered together, these data reflect the broad diversity of schools in Texas's 11 major urban school districts.

Finally, before examining the results of the regression analyses, I present descriptive statistics regarding the within-school racial achievement gaps in this sample. The table for these dependent variables can be found in Appendix A. In this paragraph, I report means of note, followed by a paragraph regarding variable ranges. Within the 1,357 campuses studied, within-school racial achievement gaps were greatest among the higher performance levels (meets and masters) on fifth grade STAAR tests. The largest mean gap was the White/African American gap at the masters level of the fifth grade math assessment, on which White students outscored their African American peers by an average of 21 percentage points. In general, gaps were narrower at the secondary, end-of-course (EOC) level. Interestingly, the masters performance level offered the widest gaps on the Algebra EOC exam, while the masters level had the narrowest gaps on the English EOC exam composite variable. (As a reminder, I aggregated English I and English II data into a single variable for ease of analysis and interpretation.) Across the sample, the lowest mean gap existed between White and Hispanic students scoring at the approaches level on the Algebra assessment, a mere 0.15 percentage points. This mean gap was unique, as it was the only one in which students of color scored higher than White students on the average, although the difference was very slight. Significantly, on each assessment and at each performance level, the White/African American gap was greater than the White/Hispanic gap by as much as 8.6 percentage points (fifth grade math, masters level). Certainly, in this initial exploration, the sample's racial achievement gap data appear consistent with the extant literature, as numerous researchers have also found that White students tend to outscore students of color on standardized tests in the United States (Aud et al., 2011; Berends, 2014; Gregory et al., 2006; Haycock, 2001; Kahlenberg, 2001; Ladson-Billings, 2010; Lee, 2002; Manning & Kovach, 2002; Mattison & Aber, 2007; Mickelson, 2001; Milner, 2010; Noguera & Wing, 2006; Orfield et al., 2008; Orfield & Lee, 2007, p. 200; Rothstein, 2004; Rubin et al., 2006; Scherff & Piazza, 2009; Williams, 2002a, 2002b; Wing, 2006).

The ranges among the racial achievement gaps variables represent an additional opportunity to conceptualize and contextualize the sample before delving into deeper analyses. Generally speaking, within this sample, wider ranges existed on reading assessments, on fifth grade assessments, at the masters performance level, and between White and African American students. The broadest range emerged within the White/African American gap on the fifth grade reading test at the meets level; at Daggett Elementary School in Fort Worth ISD, White students outscored their African American peers by a staggering 90 percentage points, while at Meadow Village Elementary School in Northside ISD, African American students scored 50 percentage points higher than White students (TEA, 2019). In comparison, the White/Hispanic gap at the masters level of the English EOC assessment offered the narrowest range, from 39 percentage points (Austin High School, Houston ISD) to -13 (Advanced Learning Academy, San Antonio ISD) (TEA, 2019). The descriptive statistics presented here suggest that racial disparities exist yet vary in the academic achievement of students in Texas's major urban school districts, reinforcing the relevance of this thesis's research questions within this research site.

Research Question 1 Results

In the first research question, I asked, "What is the nature of the relationship between magnet status (i.e., whether a school offers a magnet program or not) and within-school racial achievement gaps, controlling for other contextual variables known to influence academic outcomes?" Multiple regression offered the best strategic approach to distill this relationship, as described in Chapter Three. The model for this question can be expressed as:

$$Y_{i\ racialgap} = \beta_0 + \beta_1 X_{PWSmagnet} + \beta_2 X_{wholemagnet} + \beta_3 X_{enrollment} + \beta_{4\text{-}13} X_{districtdummy} + \beta_{14} X_{FRPLpercentage} + \beta_{15} X_{ELpercentage}$$

$$+ \beta_{16} X_{SpEdpercentage} + \beta_{17} X_{classsize} + \beta_{18} X_{advanceddegrees} + \beta_{19} X_{principalexperience} + \beta_{20} X_{teacherexperience} + u_i$$

In total, this regression took 24 forms, accounting for all possible combinations of the four assessments, the three performance levels, and the two racial dyads. I report the results of these analyses in this section, first presenting data regarding the fifth grade STAAR tests, followed by results from the high school EOC exam models.

Fifth Grade Achievement Gap Analyses

Model Fit. In the context of the fifth grade assessments, eight of the 12 models produced an *F* statistic that was significant at the $p < 0.05$ level. These eight models were the six math assessment models (addressing both racial dyads at all three performance levels on this test) and the White/Hispanic models at the meets and masters levels on the reading assessment. Within the context of this thesis, I implement Muijs's (2004) benchmarks for goodness of fit to evaluate the strength of the significant models. Muijs (2004) posits that an R^2 value of 0.11-0.30 represents a modest model fit, 0.31-0.50 a moderate fit, and greater than 0.50 a strong fit. It should be kept in mind that Muijs describes these benchmarks as "a rough guide … rule of thumb" (2004, p. 166).

Considering these guidelines, the eight statistically significant fifth grade models for the first research question can be considered as having a modest fit, as their R^2 statistics ranged from 0.115 to 0.188. At the lower end of the spectrum, the fifth grade math, approaches level, White/Hispanic model produced the R^2 value of 0.115 ($F(20, 279) = 1.80$, $p = 0.020$), indicating that this model significantly accounted for 11.5% of the variance in the within-school racial achievement gap. In contrast, the fifth grade math, masters level, White/African American gap model resulted in an R^2 value of 0.188 ($F(19, 174) = 2.12$, $p = 0.006$), significantly accounting for 18.8% of the variance in the dependent variable. The analyses related to math STAAR tests typically resulted in higher R^2 values than their corresponding reading models. Additionally, the gaps at the meets and masters levels generally produced greater R^2 values than the approaches-level models. Considered together, these findings offer interesting context for the first research question while also highlighting the modest fit of these models in terms of explaining within-school racial achievement gaps in the sample.

Model Results. Next, in examining regression coefficients, I found that neither whole school nor program-within-school magnet status were significantly associated with within-school racial achievement gaps in this sample. This was the case for all of the fifth grade models for this research question. Rather, the variables that offered coefficients significant at the $p < 0.05$ level were isolated cases of a given school district’s fixed effect, campus enrollment, or a student demographic variable impacting a particular model. In total, the exploration of the first research question in the context of fifth grade racial achievement gaps suggests that the question’s null hypothesis should not be rejected, as insufficient evidence exists to indicate that magnet status and within-school racial achievement gaps have a significant relationship, when controlling for relevant school-level variables.

End-of-Course Achievement Gap Analyses

Model Fit. Interestingly, analyses of within-school racial achievement gaps at the secondary, EOC level produced stronger R^2 values than those at the fifth grade level. Once again, eight of the 12 models produced F statistics that were significant at the $p < 0.05$ level. These eight models included three that can be considered as having a modest fit, four with a moderate fit, and one with a strong fit (Muijs, 2004). Specifically, the R^2 statistics spanned from 0.180 to 0.590, a much broader range than those of the fifth grade models. The Algebra EOC, approaches level, White/African American model produced the R^2 value of 0.180 (F(20, 117) = 1.90, p = 0.019), accounting for 18.0% of the variance in within-school racial achievement gap. The English EOC, masters level, White/African American gap model, in contrast, resulted in an R^2 value of 0.590 (F(20, 99) = 7.12, p = 0.000), significantly accounting for an impressive 59% of the variance in the dependent variable. Consistently, models of White/African American gaps offered greater R^2 values than their corresponding White/Hispanic gap models. Analyses of English EOC data also consistently produced greater R^2 statistics than corresponding Algebra analyses. For example, the English EOC, masters level, White/Hispanic gap model offered an R^2 value more than double that of the Algebra EOC, masters level, White/Hispanic gap model. The substantial R^2 values among the EOC models for the first research question starkly juxtaposed those of the fifth grade models.

Model Results. However, the EOC analyses offer similar findings to those of the fifth grade models in that magnet status, whether whole school or program-within-school, was not statistically significant at the $p < 0.05$ within any model. The only variables that offered coefficients significant at the $p < 0.05$ level were again isolated cases of individual school districts' fixed effects, campus enrollment, or student demographics influencing a given model. As with the fifth grade results, the EOC regression models fail to produce evidence that would support the rejection of the null hypothesis. Within this sample and using this methodology, there does not appear to be a significant relationship between magnet status and within-school racial achievement gaps, controlling for relevant contextual variables.

Research Question 2 Results

In the second research question, I asked if the relationship between magnet status and within-school racial achievement gaps varies by magnet theme. Again, multiple regression represented the most appropriate methodological approach. The model for this research question can be written:

$$Y_{i\ racialgap} = \beta_0 + \beta_1 X_{PWSmagnet} + \beta_2 X_{wholemagnet} + \beta_3 X_{enrollment} + \beta_{4\text{-}13} X_{districtdummy} + \beta_{14} X_{FRPLpercentage} + \beta_{15} X_{ELpercentage} + \beta_{16} X_{SpEdpercentage} + \beta_{17} X_{classsize} + \beta_{18} X_{advanceddegrees} + \beta_{19} X_{principalexperience} + \beta_{20} X_{teacherexperience} + \beta_{21\text{-}37} X_{magnetthemedummy} + u_i$$

As with the previous question, this analysis involved 24 individual models to address all possible combinations of the four assessments, the three performance levels, and the two racial dyads. Below, I report relevant results for the fifth grade STAAR tests and high school EOC exam analyses in terms of model fit and model results.

Fifth Grade Achievement Gap Analyses

Model Fit. Eight of the 12 fifth grade models produced F statistics that were significant at the $p < 0.05$ level. These models mirrored the eight contexts that were significant in the first research question's analyses (the six pertaining to fifth grade math as well as the reading, meets and masters levels, White/Hispanic models). Using Muijs's (2004) guidelines, all eight can be considered as having a modest fit, with R^2 values ranging from 0.174 to 0.231. The weakest fit occurred within the math, meets level, White/Hispanic gap model, which offered an R^2 value of 0.174 ($F(25, 274) = 2.31$, $p = 0.001$), accounting for 17.4% of the variance in the relevant achievement gap. At the opposite end of this narrow range, the math, masters level, White/African American gap produced an R^2 value of 0.231 ($F(23, 170) = 2.22$, $p = 0.002$), accounting for 23.1% of the variance in the dependent variable. Looking across these eight regression models, the White/African American gap models consistently produced greater R^2 values than corresponding White/Hispanic gap models. These characteristics of the models and their fits should be kept in mind during the subsequent discussion of statistically significant regression coefficients, particularly recalling the modest fit of each of the eight significant models.

Model Results. Across the eight significant models, six produced statistically significant regression coefficients for at least one magnet theme. Nonmagnet campuses served as the reference category for magnet theme in every regression, so each statistically significant regression coefficient for an individual theme represents a value added by that theme regarding within-school racial achievement gaps. Table 4.2 offers an at-a-glance summary of these coefficients. In the next paragraph, I outline the significant magnet theme results from the reading STAAR assessment regression analyses, followed by a paragraph regarding the math STAAR models' coefficients.

Table 4.2

Statistically Significant Fifth Grade Model Coefficients, Research Question 2

Theme	Model Detail	Coefficient	*t*	*P* > |*t*|	R^2
Fifth Grade Reading STAAR Model Results					
Fine arts	Meets, White/Hispanic	0.148	2.02	0.044	0.213
Fine arts	Masters, White/Hispanic	0.232	3.20	0.002	0.197
Leadership	Meets, White/Hispanic	-0.482	-2.67	0.008	0.213
Fifth Grade Math STAAR Model Results					
Fine arts	Approaches, White/Hispanic	0.097	2.40	0.017	0.181
Fine arts	Meets, White/Hispanic	0.178	2.68	0.008	0.174
Fine arts	Masters, White/Hispanic	0.176	2.41	0.017	0.193
Advanced academics	Approaches, White/African American	0.318	5.01	0.000	0.213
STEM/ STEAM	Approaches, White/African American	-0.111	-2.85	0.005	0.213
STEM/ STEAM	Approaches, White/Hispanic	-0.124	-2.80	0.006	0.181
STEM/ STEAM	Masters, White/Hispanic	-0.180	-2.25	0.025	0.193
Montessori	Approaches, White/African American	-0.209	-4.68	0.000	0.213
Leadership	Meets, White/Hispanic	-0.331	-2.02	0.044	0.174

The two fifth grade reading STAAR models found to have a statistically significant model fit offered a total of three magnet theme regression coefficients that were significant at the $p < 0.05$ level. In both, fine arts magnets were significantly associated with wider within-school racial achievement gaps, while one model showed that leadership magnets were associated with a narrower gap. Compared to nonmagnets, fine arts magnets were significantly associated with a wider within-school White/Hispanic gap on the fifth grade STAAR reading assessment of 14.8 percentage points at the meets level (t = 2.02, p = 0.044) and 23.2 points at the masters level (t = 3.20, p = 0.002), holding all other variables constant. In contrast, leadership magnets were significantly associated with a narrower within-school White/Hispanic gap in meets-level fifth grade STAAR reading scores of a staggering 48.2 percentage points (t = -2.67, p = 0.008) compared to nonmagnets, all else held equal. Considered together, these findings suggest that fine arts magnets may worsen White/Hispanic fifth grade reading assessment gaps, while leadership magnets may reduce this gap at the meets performance level.

Across the fifth grade math STAAR assessment models found to have significant R^2 values - all six models - I found nine significant magnet theme coefficients. Specifically, in three models, the fine arts theme was associated with wider racial achievement gaps, and one found that the advanced academics theme did so. In contrast, three models found that STEM/STEAM magnets were associated with narrower within-school racial achievement gaps, one found that Montessori magnets did so, and one found that leadership magnets did so. The details of these results are presented in the following two paragraphs.

First, fine arts and advanced academics were associated with wider fifth grade math achievement gaps within this analysis. The fine arts models, interestingly, produced significant coefficients within all three White/Hispanic fifth grade math models and no significant coefficients within the three White/African American math models. Specifically, compared to nonmagnets, fine arts magnets were associated with wider within-school White/Hispanic gaps in fifth grade STAAR math scores of 9.7 percentage points at the approaches level (t = 2.40, p = 0.017), 17.8 points at the meets level (t = 2.68, p = 0.008), and 17.6 points at the masters level (t = 2.41, p = 0.017), holding all else equal. Additionally, advanced academics magnets were associated with a wider within-school White/African American gap in approaches-level fifth grade STAAR math scores of 31.8 percentage points (t = 5.01, p = 0.000) compared to nonmagnets, holding all other variables constant. These findings suggest that both fine arts and advanced academics themes appear to exacerbate within-school racial achievement gaps on the fifth grade math STAAR, with fine arts magnets significantly impacting more of these gaps.

In comparison, STEM/STEAM, Montessori, and leadership magnets appeared to mitigate within-school racial achievement gaps on the fifth grade math STAAR assessment. STEM/STEAM magnets were associated with a wider within-school White/African American gap in fifth grade STAAR math scores of 11.1 percentage points at the approaches level (t = -2.85, p = 0.005), compared to nonmagnets and holding all else equal. This theme was also associated with narrower within-school White/Hispanic gaps on the math assessment of 12.4 points at the approaches level (t = -2.80, p = 0.006) and 18 points at the masters level (t = -2.25, p = 0.025) compared nonmagnet schools and all else held equal. In addition, compared to nonmagnets, Montessori magnets were significantly associated with a narrower within-school approaches-level White/African American gap on the fifth grade math STAAR of 20.9 percentage points (t = -4.68, p = 0.000), holding all other variables constant. Leadership magnets were associated with a narrower within-school White/Hispanic gap at the meets level of the fifth grade math STAAR of 33.1 percentage points ($t = -2.02$, $p = 0.044$) compared to nonmagnets, all else held equal. These results suggest that STEM/STEAM, Montessori, and leadership magnet themes lessen fifth grade math achievement gaps.

Considered together, these findings suggest that magnet themes have significant and varied relationships with within-school racial achievement gaps on the fifth grade math STAAR assessments. Six of the significant coefficients related to White/Hispanic gaps, implying that these gaps may be particularly sensitive to magnet themes. Most notably, fine arts magnets appeared to exacerbate within-school racial achievement gaps on the math assessment, and STEM/STEAM magnets appeared to lessen them. The models related to advanced academics, Montessori, and leadership magnets each produced one significant coefficient, suggesting that these magnets significantly impact racial achievement gaps but within narrower contexts than fine arts and STEM/STEAM magnets. Considering these findings and those presented in the previous paragraph, the null hypothesis for this research question should be rejected, as magnet theme appears to have a significant relationship with within-school racial achievement gaps within this sample, controlling for contextual variables.

After reviewing these findings, it should be noted that magnet status (neither whole school nor program-within-school) was not found to be significantly associated with fifth grade within-school racial achievement gaps in the first research question. This affirms my earlier theory that examining the effects of magnet programs without disaggregating the impacts of individual themes can mask significant heterogeneity across the broad field of magnet programming. Based upon the results presented in this section, it appears that certain magnet themes offer beneficial influences on students' opportunities to learn, while others have problematic impacts.

Lastly, across the eight significant fifth grade models for the second research question, variables other than magnet themes offered significant regression coefficients in isolated scenarios. These included specific districts' fixed effects, campus enrollment, and student demographics. Each of these findings will not be specifically reported in this chapter, given their lack of direct alignment to this thesis's research questions and their lack of consistency across models, suggesting limited influence. I will address this further at the end of this chapter, however.

End-of-Course Achievement Gap Analyses

Model Fit. Six of the 12 EOC models had a statistically significant model fit, as indicated by their *F* statistics. Five of these occurred within the English EOC context, including each of the three White/Hispanic gap models as well as the meets- and masters-level White/African American gap models. Additionally, the Algebra I EOC, meets level, White/African American gap produced a significant *F* statistic. Five of these six significant models can be described as having a moderate fit (Muijs, 2004), while the sixth - English EOC, masters level, White/African American gap - had a strong fit. This sixth model offered an R^2 value of 0.664 ($F(29, 90) = 5.62$, $p = 0.000$), accounting for an impressive 66.4% of the variance in the dependent variable. In contrast, the English EOC, meets level, White/Hispanic model had the lowest R^2 statistic, 0.320, ($F(30, 103) = 1.62$, $p = 0.040$), accounting for 32% of the variance in the relevant achievement gap. It should be noted that even this weakest significant model can still be characterized as having a moderate fit (Muijs, 2004). Examining the trends across the five significant English EOC models, White/African American gap models consistently produced greater R^2 values than their corresponding White/Hispanic models, and model fit increased as the performance levels increased, such that masters-level models offered the greatest R^2 statistics. As with the fifth grade models, these model fit results and trends provide meaningful context for the following coefficient-level results.

Model Results. Five of the six significant EOC models for this research question produced significant coefficients for magnet themes. Altogether, these included five significant coefficients from English EOC analyses and one from the significant Algebra EOC model. Table 4.3 summarizes these findings. I describe the results in the following paragraphs, first those from the English EOC models, then those from the Algebra model.

Table 4.3

Statistically Significant EOC Model Coefficients, Research Question 2

Theme	Model Detail	Coefficient	*t*	$P > \|t\|$	R^2
English EOC Model Results					
Language	Masters, White/African American	0.216	2.55	0.013	0.644
Language	Meets, White/Hispanic	0.238	2.50	0.014	0.320
Language	Masters, White/Hispanic	0.206	3.01	0.003	0.486
Montessori	Masters, White/Hispanic	0.283	2.96	0.004	0.486
Early college	Meets, White/African American	-0.248	-3.42	0.001	0.458
Algebra EOC Model Results					
Early college	Meets, White/African American	-0.247	-2.97	0.004	0.440

First, the five significant English EOC models offered a total of five significant magnet theme coefficients, all at the meets or masters level. Three of these coefficients indicated that language magnets related to wider within-school racial achievement gaps compared to nonmagnet schools, one suggested this for Montessori magnets, and a final indicated that early college magnets were associated with a narrower achievement gap. First, compared to nonmagnets, language magnets were found to associate with wider within-school White/Hispanic gaps on English EOC assessments of 23.8 percentage points at the meets level (t = 2.50, p = 0.014) and 20.6 points at the masters level (t = 3.01, p = 0.003), holding all other variables constant. Language magnets were also associated with a wider masters-level White/African American gap on these assessments of 21.6 percentage points (t = 2.55, p = 0.013), compared to nonmagnets and holding all else equal. Additionally, Montessori magnets were associated with a wider White/Hispanic gap at the masters level of the English assessments of 28.3 points (t = 2.96, p = 0.004) compared to nonmagnets, holding all other variables constant. In contrast, early college magnets were significantly associated with a narrower within-school White/African American gap on these assessments of 24.8 percentage points at the meets level (t = -2.47, p = 0.001) compared to nonmagnets, all else held equal. Considered together, these results suggest that language as well as Montessori magnets may be detrimental in addressing English EOC racial achievement gaps, particularly at the meets and masters level, while early college magnets may be beneficial in reducing the meets-level White/African American gap on English EOC assessments.

The Algebra EOC regression analyses offered one significant magnet theme coefficient. Compared to nonmagnets, early college magnets were significantly associated with a narrower within-school White/African American gap on the Algebra EOC assessment of 24.7 percentage points at the meets level (t = -2.97, p = 0.004), holding all other variables constant. The lack of additional significant coefficients should not be surprising, as only one Algebra EOC model could be considered due to the fact that the other five were not statistically significant at the $p < 0.05$ level. Nevertheless, this finding as well as those presented in the paragraph regarding English EOC analyses suggest that the null hypothesis for this research question should be rejected in the EOC context, as magnet theme appears to have a significant relationship with within-school racial achievement gaps at the secondary school level as well, controlling for contextual variables.

Additionally, as with the fifth grade regression models, the EOC models identified some significant correlation coefficients related to individual school districts' fixed effects, campus enrollment, and certain student demographic variables. Class size and the percentage of teachers with advanced degrees were also significantly associated with select changes in racial achievement gaps in select models. However, again, I chose to focus reporting on the independent variables directly addressed in this research question - magnet theme - rather than individual cases of control variables. I will address this choice further at the end of this chapter.

Significant Control Variable Coefficients

As noted earlier in this chapter, several control variables produced statistically significant coefficients in this analysis. In this section, I will briefly address these coefficients. I summarize control variable coefficients of note within the analysis of the second research question as this context addressed the research purpose more directly and offered greater insight into magnets' relationship with achievement gaps. Additionally, I report these findings in tabular format in Appendix B.

I incorporated a total of 20 control variables in the model for the second research question. 11 of these variables produced statistically significant regression coefficients at least once within the 14 models with F statistics significant at the $p < 0.05$ level. Increases in the percentage of students who receive free or reduced-price lunch, the percentage of English learners, and average class size were significantly associated with two narrowing gaps. On the other hand, increases in a school's percentage of students who receive special education services or the percentage of teachers with advanced degrees were associated with wider achievement gaps. The number of students enrolled were associated with one narrower gap and two wider gaps. Similarly, the school district control variables produced few significant coefficients in most cases. These district coefficients proved quite varied, with some district fixed effects associating with wider achievement gaps while others related to narrower gaps. Considered together, these results do not suggest a broad, consistent pattern of influence of any of these control variables on within-school racial achievement gaps.

I have chosen to not delve deeper into these results for several reasons. First, these findings do not align with either of my research questions, in which I sought to explore the relationship between magnet status or theme and within-school racial achievement gaps. This relationship has historically represented a gap in the literature (Betts et al., 2015; Bifulco et al., 2009; Gamoran, 1996; Wang et al., 2014, 2017; Wang & Herman, 2017), contributing to the significance of this study. The extant literature offers a great deal of preexisting insight into the impact of these control variables on achievement gaps.

So, I wish to center my analysis on what makes my thesis unique in terms of addressing an important gap in the extant literature. Finally, the control variable findings within the context of my study do not offer any broad patterns, only relevance in case-by-case scenarios. My research purpose is for the reader to garner new and impactful knowledge regarding the influence of individual magnet themes on within-school racial achievement gaps, so I have kept this as the focus of my analysis. However, I provide additional detail regarding these results in Appendix B.

Conclusion

The primary takeaway from this analysis is that, within Texas's 11 major urban school districts, certain magnet themes appeared to contribute to within-school racial achievement gaps while others appear to lessen them in the 2018-2019 school year. This represents a meaningful overall finding, as descriptive statistics demonstrated that achievement gaps were rampant across the units of analysis in this work, with White students' test score averages almost always higher than those of their Hispanic and African American peers.

14 of the 24 models stemming from the key second research question proved to be significantly significant at the $p < 0.05$ level, including every fifth grade math STAAR assessment model and all but one English EOC model. Across these 14 models, I identified a total of 18 magnet theme coefficients that were significantly associated with a within-school racial achievement gap at the $p < 0.05$ level. Examining the 18 significant magnet theme findings, elementary fine arts magnets and secondary language magnets appear to relate to multiple widening within-school racial achievement gaps, compared to schools without magnet programs. In comparison, elementary STEM/STEAM,

elementary leadership, and secondary early college magnets appear to be associated with multiple narrowing gaps. Of the 17 magnet themes included in this sample, 10 were not found to be significantly associated with an achievement gap in any model. These were (in order of prevalence): CCR, gifted and talented/Vanguard, IB/international, single-gender, health, humanities, business, law, PE, and literature. The mechanisms behind these relationships between magnet themes racial and achievement gaps, however, are outside of the scope of this study. In the “Implications for Researchers” section of Chapter Five, I present my personal theories regarding potential factors behind each of these findings.

Lastly, it is worth noting two patterns that emerged across the 18 significant magnet theme coefficients. 12 of these instances occurred within White/Hispanic achievement gap models, and 13 presented in models at the meets or masters performance levels. These findings suggest that magnet programming may be particularly influential for Hispanic students as well as for improving Hispanic and African American students’ performance at the higher passing levels on STAAR exams. The latter observation is particularly relevant as scoring at the meets or masters level indicates that a student has acquired the key content of their course and can open doors to advanced academic tracks in secondary education, compared to passing at merely the approaches level. These patterns provide an interesting opportunity for further exploration in future research.

Considered together, these results reinforce important findings from existing research that demonstrate how the opportunities schools and districts provide for their students, including magnet programming, impact student learning, including in inequitable ways. Expanding upon a quote from Ballou (2007), my findings indicate that "in at least some times and places, students benefit from enrolling in magnet schools" (p. 31) offering certain - but not all - magnet themes.

Chapter V

Conclusion

In Chapter One of this thesis, I described the purpose for this study - filling a gap in the literature regarding the effects of individual magnet themes on within-school racial achievement gaps in an effort to identify best practices for mitigating opportunity gaps. The results of my analyses, as reported in the previous chapter, revealed magnet program themes that were associated with wider or narrower within-school achievement gaps within the sample, fulfilling the first aspect of this research purpose. In this chapter, I address the second half of my thesis's purpose by outlining implications for researchers, practitioners, and policymakers. My research implications also address the mechanisms that I believe may be influencing my findings. I also address the limitations within this work.

Implications for Researchers

Reflecting upon the results reported in Chapter Four, I have identified a number of implications for the education research community. First and foremost, magnet researchers must let go of the tradition of misrepresenting magnet schools as a homogeneous category. I believe the most important finding of my thesis is that this methodology offered zero significant findings at the $p < 0.05$ level for the first research question yet numerous for the second. The lack of findings among the models that included only magnet status, not theme, indicated that neither whole school nor program-within-school magnets were associated with any changes in within-school racial achievement gaps compared to nonmagnet schools. Put another way, magnet schools, as a collective, do not appear to have a significantly different impact upon within-school racial achievement gaps than do nonmagnets. The disaggregated analyses of the second research question, however, unveiled a great deal of heterogeneity across magnet themes that was masked within the first set of models. It should be noted that I am not alone in calling for recognizing and exploring differences in magnet themes within the context of education research (Betts et al., 2015; Bifulco et al., 2009; Gamoran, 1996; Wang et al., 2014, 2017; Wang & Herman, 2017).

Additionally, a number of future research avenues have become apparent. The methodology of this thesis could be replicated with different and broader samples to explore the relationships of magnet themes and within-school racial achievement gaps in other contexts. For example, the results of this study cannot fully generalize to other states, so exploring these relationships in other areas could prove beneficial. A national sample with a larger dataset could also increase the statistical power of the models, especially for variables representing less common magnet themes. Less common magnets in this sample, such as law magnets, may have impacts on achievement gaps yet did not meet the threshold for statistical significance in this analysis due to their minimal representation in the sample. Additionally, this methodology could be replicated with other racial dyads besides White/Hispanic and White/African American given a sample that sufficiently represented other racial groups. A similar study incorporating controls for schools' percentages of students of specific racial groups could provide an interesting comparison point to the findings of this thesis. A benefit of the methodology of this thesis is that its general structure could offer a scaffold on which to build further explorations of the relationship between magnet themes and within-school racial achievement gaps.

A need exists for investigation of the specific programmatic factors that influence the relationships between magnet themes and within-school racial achievement gaps as well. This work identified benefits and disadvantages of specific magnet themes, but it was outside of its scope to identify the mechanisms by which these programs have impacted achievement gaps. To further benefit the education field, future research into these mechanisms is warranted and necessary.

I speculate that STEM/STEAM magnet programs may have narrowed achievement gaps on the fifth grade math STAAR by increasing teaching time, resources, teacher training, and school investment in math education. Perhaps the association between elementary leadership magnets and narrowing White/Hispanic reading and math achievement gaps stemmed from students' increased voice in their schools, which could relate to more speaking opportunities, benefitting English learners.

I believe the association between early college magnets and narrowing gaps on the English and Algebra EOC exams may be strongly influenced by selection bias. Selection bias, also called sampling bias, can occur when samples are potentially not representative of the wider population and may consist of a disproportionate number of individuals with certain relevant characteristics (American Psychological Association, 2020). School choice inherently introduces this factor in magnet research. Specifically, early college magnets may overrepresent students who have exemplary academic records, are intrinsically motivated, have strong support systems, and/or had middle school guidance counselors that recognized their potential. Additionally, early college students are by definition on advanced academic tracks, which can render the foundational English and Algebra assessment content less challenging for these students.

On the other hand, I believe the widening reading and math achievement gaps at elementary fine arts magnets likely resulted from these schools taking time away from core academic subjects (time which may be especially necessary for students who did not have as many academic advantages before or during elementary school) to teach the arts. This logic could also extend to the widening achievement gaps at language magnets.

Further, two of the three widening gaps at language magnets were between White students and Hispanic students, who may be English learners taking on a third language beyond English and Spanish, and trilingualism is no easy feat.

Each of these speculations are based upon my experiences as a magnet educator in one of the districts within this research sample. However, personal perspectives based on anecdotes cannot replace empirical data. Further research into the mechanisms driving the relationships I have found in this research is necessary to improve guidance for schools and districts. The OTL framework may be beneficial for this work, as an explicit exploration of its four components - content coverage, content exposure, content emphasis, and quality of instructional delivery (Stevens, 1993, 1996; Wang, 1998) - could further distill how certain magnet themes are influencing opportunity and achievement gaps. Additionally, research addressing the mechanisms I theorize would likely benefit from a mixed methods approach to capture multiple facets of the interplay between learning opportunities and student outcomes in magnet schools. The research avenues and methodological adaptations I have outlined here could expand the field's knowledge base in terms of *how* magnet programs impact within-school racial achievement gaps at magnet and nonmagnet schools. With this knowledge, scholars could further identify strategies for schools and districts to break the pernicious cycle of racial opportunity gaps in American public education.

Implications for Practitioners

In this section, I offer implications for school and district leaders who develop and implement magnet programs. My findings offer several key takeaways regarding theme selection for both elementary and secondary schools. Additionally, patterns across the racial dyads in the regression models offer implications for magnet programming as well.

First, the findings for my second research question suggest certain magnet themes may relate to within-school racial achievement gaps in Texas's 11 major urban school districts. Specifically, and in comparison to nonmagnet schools, STEM/STEAM magnets and leadership magnets are associated with narrower within-school racial achievement gaps on fifth grade reading and math STAAR assessments. On the other hand, fine arts magnets were consistently associated with wider gaps on these assessments. A single significant regression coefficient suggested that Montessori magnets relate to a narrower fifth grade math, approaches level, White/African American gap, while a single model indicated that advanced academic magnets relate to a wider gap.

It should be noted that I found no other magnet themes to be significantly associated with elementary schools' racial achievement gaps. Each of these results directly implies which magnet themes school and district leaders - particularly those serving diverse student populations - should consider for implementation and further examine for best practices. Primarily, it appears that STEM/STEAM magnets and leadership magnets should be continued or adopted. Both caution and further evaluation are warranted when districts wish to develop or maintain fine arts magnets in order to determine their potential impacts on gaps in opportunity and achievement in the specific context. For example, practitioners could conduct equity audits (Green, 2017) to evaluate existing fine arts magnets and identify strengths and weaknesses regarding the learning opportunities these programs provide. Certainly, such evaluations would be highly beneficial within the context of magnet programs of any theme, but my findings suggest that elementary fine arts magnets in particular should be implemented with caution and close consideration.

Similarly, the results of the second research question also imply magnet themes that may influence within-school racial achievement gaps at secondary schools in this sample. It should be noted that the model fits (in terms of proportion of overall variance explained) of the EOC-level regressions for this research question tended to be greater than those of the fifth grade analyses, suggesting that magnet themes may impact secondary schooling more intensely, contributing to the importance of equitable magnet programming at this level. This could reflect an increased exposure time to the "treatment" of magnet schools among high school students who may have attended magnet middle and/or elementary schools. Primarily, these findings showed that early college magnets were associated with narrower within-school racial achievement gaps on English and Algebra EOC assessments, compared to nonmagnet schools. Language magnets, on the contrary, related to wider English EOC racial achievement gaps. This methodology did not lead to any significant findings regarding the other magnet themes offered in Texas's major urban school districts. As with the elementary-level results, these suggest that school and district leaders who wish to lessen racial gaps should maintain and expand the use of early college magnets while exercising caution and thoughtful evaluation when considering continuing or introducing secondary language magnets.

Lastly, it should be noted that two-thirds of the significant findings for the second research question occurred within the context of White/Hispanic achievement gaps. This pattern suggests that magnet themes may be especially impactful for shaping the learning experiences of Hispanic students. Hispanic students represent both the largest population and the fastest growing population in Texas public schools (TEA, 2020, August). So, implementing school programming that improves their learning opportunities and academic achievement is particularly important. Practitioners who serve Hispanic students should be cognizant of this as they select magnet themes, given the finding that magnet themes offer a particular opportunity to influence the educational experiences of this important student population.

Implications for Policymakers

In examining the results of this thesis, I have also recognized implications for policymakers. I outline policy implications at the local, state, and federal levels in this section.

Policy implications for local district leaders are closely tied to the recommendations for practitioners I outlined above. Primarily, district policymakers should support an increase in the implementation of elementary STEM/STEAM magnets, elementary leadership magnets, and early colleges in their districts, while exercising caution with elementary fine arts and secondary language magnets. District leaders could also recommend replacing fine arts or language magnets with STEM/STEAM, leadership, or language programs on campuses with wide racial achievement gaps. Local evaluation, such as equity adults (Green, 2017), can support informed decisions in this regard. Funding policies can be implemented to incentivize the use of research-supported magnet themes. Additionally, policymakers should ensure that stakeholders are aware of which magnet themes are likely to benefit the community's children, especially when discussing the adoption, expansion, or closure of a magnet. District policymakers are often empowered with a great deal of authority in terms of magnet programming, as they shape district finances and school choice policies, so these leaders should advocate for magnet themes that will benefit the students they serve.

State policymakers, including those at the Texas Education Agency, should increase political and financial support for best practices in magnet programming, particularly given the state's trend of championing school choice. It may be appropriate for the state to supplement funding for magnet programs with research-supported themes. In addition, Domain III of the state's three-part accountability system directly measures the presence and changes of within-school and -district achievement gaps. The TEA could incorporate magnet programming guidance in the support they provide to districts and campuses with poor Domain III scores, informing them of magnet themes that could narrow the specific gaps they are experiencing. Based upon their extensive policymaking power and using their rich data warehouse, state policymakers can identify and support targeted contexts where magnet programming may address racial achievement gaps.

Lastly, my results have implications for federal policymakers. I found that certain magnet themes are indeed associated with narrowing achievement gaps, justifying the continuation of the Magnet Schools Assistance Program described in Chapter Two. The government's financial investment in this grant program, however, has been stagnant and uncontrolled for inflation since its inception in 1985 (Frankenberg et al., 2008; Rossell, 2005). As opportunity gaps and achievement gaps continue to impact students' lives - particularly those of the consistently growing proportion of students of color - increased financial support is warranted. I recommend that federal policymakers invest more funds in the MSAP, including incentivizing the implementation of research-supported magnet themes with supplemental funding. Among their guidance to grantee districts, the MSAP can and should inform grant recipients of which magnet themes address racial achievement gaps. Lastly, the U.S. Department of Education collects extensive data from every public school in the nation each year, including detailed data regarding students, employees, test scores, finances, disciplinary events, and basic characteristics such as grades served. I call for the addition of two new labels to benefit future magnet researchers - a magnet/nonmagnet tag as well as a magnet theme field for schools that identify as magnets. With schools labeled accordingly, researchers will be able to conduct rigorous, national-level quantitative studies exploring the impacts of magnet programs of various themes. Such national studies are currently immensely challenging due to an inability to easily sort campuses by magnet status and theme. By increasing magnet funding and paving the way for future magnet research, federal policymakers can support the implementation and continued exploration of magnet program practices that improve outcomes for our country's children.

Limitations

Finally, a number of limitations exist in this work, as with any, that should be noted. In particular, limitations exist in terms of the research design, the dependent variable operationalization, and the sample.

This thesis involved a school-level research design. An inherent limitation of this approach, in comparison to one in which students are the unit of analysis, is the inability to control for student-level variables which could increase the potential for calculating a more precise value-added effect. For example, controlling for students' length of enrollment at their campus or previous academic achievement was not possible within this design. Additionally, quantitative research can fail to capture the richness of students' lived experiences, including their experiences with opportunity gaps. The methodological choice to conduct a school-level quantitative study certainly does not invalidate my thesis's findings, but additional approaches, such as student-level and qualitative work, could further illuminate the nexus between magnet programming and students' opportunities to learn. This logic extends to teacher-level data as well, considering the significant impact that teachers have on student outcomes (Hattie, 2009).

Next, the use of standardized test scores as the basis for the dependent variable offers limitations as well. Each assessment compresses the full complexity of students' educational experiences into a single datapoint from one test on one day. Additionally, as described in Chapter Two, standardized tests do not necessarily align with all magnet programs' missions (Blank & Archbald, 1992; Siegel-Hawley & Frankenberg, 2013), as "many magnet school programs are not specifically aimed at building the skills reflected on standardized tests" (Blank & Archbald, 1992, p. 87). As described in Chapter One, however, my research purpose does align with students' mastery of key academic content areas, and standardized test data represent the industry standard for measuring this.

Limitations exist within my sample and its available data as well. Again, a number of magnet themes were present but uncommon within the 11 Texas major urban school districts, hindering their ability to meet the threshold of statistical significance. Several key input variables regarding the implementation of magnet programs were unavailable to me, such as the percentage of students enrolled in schools offering program-within-school that actually participated in the magnet. Magnets' admissions policies were also inaccessible, precluding their inclusion as a potential control variable. Lastly, there likely exist additional input variables - beyond class size, teachers' advanced degrees, principal experience, and teacher experience - that could help explain the relationship between magnet programming and achievement gaps. Magnet program funding, the percentages of instructional time allocated for traditional academic subjects vs. magnet curricula, teacher turnover, district policies regarding how students are assigned to specific magnet programs within a district represent four such examples. The lens of the OTL framework calls for a deeper exploration of these inputs. Future research, as described above, with access to additional data could mitigate these limitations within the sample.

Lastly, there exists a nonzero possibility of researcher bias that should be noted. I have served as a magnet school educator for most of my career, so I entered this work with experiences, perspectives, and opinions regarding magnet programs. However, a quantitative methodology reduces the potential for my experiences to influence my results, and no incidences of this come to mind.

Conclusion

As an educator, there have been times when I have felt utterly powerless against the systemic trends of racial opportunity and achievement gaps in our field and our country. The results of this thesis, however, indicate that school and district leaders have a potentially powerful tool at their disposal to combat these gaps - specific magnet themes. My regression analyses revealed that, within Texas's major urban school districts, specific magnet themes related to narrowing within-school racial achievement gaps by as many as 48.2 percentage points (leadership magnets, fifth grade reading, meets level, White/Hispanic gap). In particular, elementary STEM/STEAM, elementary leadership, and secondary early college magnets were associated with narrower racial achievement gaps, while elementary fine arts and secondary language magnets were associated with wider gaps. Much remains to be learned regarding magnet programming and how it can be leveraged to equitably enhance students' opportunities to learn, but one strategy appears beneficial - the adoption of specific magnet themes with an evidenced record of decreasing within-school racial achievement gaps.

References

A snapshot of magnet schools in America. (2017). Magnet Schools of America. http://magnet.edu/files/Snapshot-of-Magnet-Schools-in-America.pdf

Achen, C. (1982). *Interpreting and using regression*. SAGE Publications, Inc.

Adcock, E. P. & Phillips, G. W. (2000, April 24). *Accountability evaluation of magnet school programs: A value-added model approach*. Annual Meeting of the American Educational Research Association, New Orleans, LA.

Alemán, E. (2006). Is Robin Hood the "prince of thieves" or a pathway to equity?: Applying critical race theory to school finance political discourse. *Educational Policy*, *20*(1), 113–142.

Alexander v. Holmes County Board of Education, 396 US 19 (1969).

Allensworth, E. M. & Rosenkranz, T. (2000). *Access to magnet schools in Chicago*. Consortium on Chicago School Research.

Allport, F. H., Allport, G. W., Babcock, C., Bernard, V. W., Bruner, J. S., Cantril, H., Chein, I., Clark, K. B., Clark, M. P., Cook, S. W., Dai, B., Davis, A., Frenkel-Brunswik, E., Gist, N. P., Katz, D., Klineberg, O., Krech, D., Lee, A. M., MavIver, R. M., … Williams, R. M. (1953). The effects of segregation and the consequences of desegregation: A social science statement. Appendix to appellant's brief: Brown v. Board of Education of Topeka. *Minnesota Law Review*, *37*(6), 427–439.

American Psychological Association. (2016). *Effects of poverty, hunger, and homelessness on children and youth*. http://www.apa.org/pi/families/poverty.aspx

American Psychological Association. (2020). Sampling bias. In *American Psychological*

Association dictionary of psychology. https://dictionary.apa.org/sampling-bias

Amos v. Board of School Directors of City of Milwaukee, 408 F. Supp. 765 (1976).

Aratani, Y. (2009). *Homeless children and youth causes and consequences*. National Center for Children in Poverty, Columbia University Mailman School of Public Health. http://www.nccp.org/publications/pdf/text_888.pdf

Archbald, D. A. (2004). School choice, magnet schools, and the liberation model: An empirical study. *Sociology of Education, 77*, 283–310.

Arcia, E. (2006). Comparison of the enrollment percentages of magnet and non-magnet schools in a large urban school district. *Education Policy Analysis Archives*, *14*(33).

Armstrong v. Board of School Directors of the City of Milwaukee, 408 F. Supp. 765 (1976).

Arthur v. Nyquist, 415 F. Supp. 904 (1976).

Aud, S., Hussar, W., Kena, G., Bianco, K., Frohlich, L., Kemp, J., Tahan, K., Mallory, K., Nachazel, T., & Hannes, G. (2011). *The condition of education 2011*. U.S. Department of Education, National Center for Education Statistics. http://files.eric.ed.gov/fulltext/ED520001.pdf

Austin-Archil, M.L. (2019). *Language teachers' perceptions of barriers to new language acquisition for English language learners.* (Publication No. 27735660) [Doctoral dissertation, Walden University]. ProQuest Dissertations Publishing.

Avila, K. (2015). The call to collaborate: Key considerations as ELD and classroom teachers begin to align new standards. *ORTESOL Journal*, *32*(1), 33-43. https://ortesol.wildapricot.org/Publications

Bailey, M.A. (2015) *Real stats: Using econometrics for political science and public policy*. Oxford University Press.

Ballou, D. (2007). Magnet schools and peers: Effects on student achievement. *Unpublished*.

Ballou, D. (2009). Magnet school outcomes. In M. Berends, M. Springer, D. Ballou, & H. Walberg (Eds.), *Handbook of research on school choice* (pp. 409–426). Routledge.

Ballou, D., Goldring, E., & Liu, K. (2006). *Magnet schools and student achievement*. National Center for the Study of Privatization in Education at Teachers College, Columbia University.

Bank, B. J. & Spencer, D. A. (1997, March 24). *Effects of magnet programs on educational achievement and aspirations*. Annual Meeting of the American Educational Research Association, Chicago, IL.

Barnett, A.T. (2020). *Mississippi teachers who are nationally board certified or those with advanced degrees: Effects on student achievement* (1803). [Doctoral dissertation, University of Mississippi]. Electronic Theses and Dissertations.

Bauch, P. A. & Goldring, E. B. (1995). Parent involvement and community responsiveness: Facilitating the home-school connection in schools of choice. *Educational Evaluation and Policy Analysis*, *17*(1), 1–21.

Bauch, P. A. & Goldring, E. B. (1996). Parent involvement and teacher decision making in urban high schools of choice. *Urban Education*, *31*(4), 403–431.

Berends, M. (2014). The evolving landscape of school choice in the United States ...... In H. R. Milner IV, & K. Lomotey (Eds.). In H. R. Milner & K. Lomotey (Eds.), *Handbook of Urban Education* (pp. 451–473). Routledge.

Betts, J. R., Kitmitto, S., Levin, J., Bos, J., & Eaton, M. (2015). *What happens when schools become magnet schools?: A longitudinal study of diversity and achievement*. American Institutes for Research, U.S. Department of Education.

Betts, J. R., Rice, L. A., Zau, A. C., Tang, Y. E., & Koedel, C. R. (2006). *Does school choice work?: Effects on student integration and achievement*. Public Policy Institute of California.

Bifulco, R., Cobb, C. D., & Bell, C. (2009). Can interdistrict choice boost student achievement?: The case of Connecticut's interdistrict magnet school program. *Educational Evaluation and Policy Analysis*, *31*(4), 323–345.

Blank, R. K. (1989). *Educational effects of magnet high schools*. Conference on Choice and Control in American Education, Madison, WI.

Blank, R. K. & Archbald, D. A. (1992). Magnet schools and issues of education quality. *Clearing House*, *66*(2), 81–87.

Blank, R. K., Dentler, R. A., Baltzell, D. C., & Chabotar, K. (1983). *Survey of magnet schools: Analyzing a model for quality integrated education* [Final Report of a National Study for the U.S. Department of Education]. James H. Lowry and Associates.

Blank, R. K., Levine, R. E., & Steele, L. (1996). After 15 years: Magnet schools in urban education. In B. Fuller, R. F. Elmore, & G. Orfield (Eds.), *Who chooses? Who loses? Culture, institutions, and unequal effects of school choice* (pp. 154–172). Teachers College Press.

Board of Education of Oklahoma City Public Schools v. Dowell, 498 US 237 (1991).

Brown v. Board of Education of Topeka I, 347 US 483 (1954).

Brown v. Board of Education of Topeka II, 349 US 294 (1955).

Buckner, J. C., Beardslee, W. R., & Bassuk, E. L. (2004). Exposure to violence and low-income children's mental health: Direct, moderated, and mediated relations. *American Journal of Orthopsychiatry, 74*(4), 413-423.

Campbell, N.D. & Lopez, E.J. (2008). Paying teachers to earn advanced degrees: Evidence on student performance in Georgia. *The Journal of Private Enterprise, 24*(1), 33-49.

Capacchione v. Charlotte-Mecklenburg Board of Education, 57 F. Supp 2d 228 (1999).

Chang, M., Jorrín Abellán, I.M., Wright, J., Kim, J., & Gaines, R.E. (2020). Do advanced degrees matter?: A multiphase mixed-methods study to examine teachers' obtainment of advanced degrees and the impact on student and school growth. *Georgia Educational Researcher, 1*(5), 62-89.

Christenson, B., Eaton, M., Garet, M. S., Miller, L. C., Hikawa, H., & DuBois, P. (2003). *Evaluation of the Magnet Schools Assistance Program, 1998 grantees* (2003-15). U.S. Department of Education, Office of the Under Secretary, Policy and Program Studies Service. http://www2.ed.gov/rschstat/eval/choice/magneteval/finalreport.pdf

Chubb, J. E. & Moe, T. M. (1990). *Politics, markets, and America's schools*. The Brookings Institution.

Clewell, B. C. & Joy, M. F. (1990). *Choice in Montclair, New Jersey: A policy information paper*. ETS Policy Information Center, Education Testing Service.

Clinchy, E. (1993). Magnet schools matter: Why are we ignoring our one, sure, proven path to successful reform? *Education Week*, *13*(4), 28, 31.

Clotfelter, C. T. (2004). *After Brown: The rise and retreat of school desegregation*. Princeton University Press.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). Federal oversight, local control, and the specter of “resegregation” in Southern schools. *American Law and Economics Review*, *8*(2), 347–389.

Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education* (7th ed.). Routledge.

Crain, R. L., Heebner, A. L., & Si, Y.-P. (1992). *The effectiveness of New York City’s career magnet schools: An evaluation of ninth grade performance using an experimental design*. National Center for Research in Vocational Education.

Cullen, J. B., Jacob, B. A., & Levitt, S. (2005). *The effect of school choice on student outcomes: Evidence from randomized lotteries* (Working Paper No. 10113). National Bureau of Economic Research.

Darling-Hammond, L. (1997). *The right to learn: A blueprint for creating schools that work*. Jossey-Bass Publishers.

Davis, T. M. (2014). School choice and segregation: “Tracking” racial equity in magnet school. *Education and Urban Society*, *46*(4), 399–433.

Dentler, R. A. (1990). Conclusions from a national study. In N. Estes, D. U. Levine, & D. R. Waldrip (Eds.), *Magnet schools: Recent developments and perspectives* (pp. 59–83). Morgan Printing and Publishing, Inc.

Deo, S. R. (2008). Where have all the Lovings gone: The continuing relevance of the movement for a multiracial category and racial classification after Parents

Involved in Community Schools v. Seattle School District No. 1. *Journal of Gender, Race & Justice*, *11*(3), 409–452.

Dowd, A. J., Friedlander, E., & Guajardo, J. (2014). Opportunity to learn (OTL): A framework for supporting learning? In N. V. Varghese (Ed.), *From schooling to learning* (pp. 95–109). UNESCO, International Institute for Educational Planning.

Doyle, D. P. & Levine, M. (1984). Magnet schools: Choice and quality in public education. *The Phi Delta Kappan*, *66*(4), 265–270.

Eckes, S.E. & Swando, J. (2009). Special education subgroups under NCLB: Issues to consider. *Teachers College Record, 111*(11), 2479-2504.

Eisenberg v. Montgomery County Public Schools, 19 F. Supp. 2d 449 (1998).

Elementary and Secondary Education Act, 89–10 Pub.L. (1965).

Elliott, W. (2013). The effects of economic instability on children's educational outcomes. *Children and Youth Services Review, 35*(3), 461-471.

Engberg, J., Epple, D., Imbrogno, J., Sieg, H., & Zimmer, R. (2011). *Bounding the treatment effects of education programs that have lotteried admission and selective attrition*. National Center for the Study of Privatization in Education, Teachers College at Columbia University.

Fleming, N. (2012, May 9). Magnets adjust to new climate of school choice: Longtime strategy now in the mix with charters. *Education Week*, *31*(30), 1–17.

Frankenberg, E. (2008). School segregation, desegregation, and integration: What do these terms mean in a post-Parents Involved in Community Schools, racially transitioning society? *Seattle Journal for Social Justice*, *5*(2), 553–590.

Frankenberg, E. & Le, C. Q. (2008). The post-parents involved challenge: Confronting extralegal obstacles to integration. *Ohio State Law Journal*, *69*, 1015–1072.

Frankenberg, E. & Lee, C. (2002). *Race in American public schools: Rapidly resegregating school districts*. The Civil Rights Project at Harvard University.

Frankenberg, E. & Siegel-Hawley, G. (2010). Choosing diversity: School choice and racial integration in the age of Obama. *Stanford Journal of Civil Rights & Civil Liberties*, *6*(2), 219–252.

Frankenberg, E., Siegel-Hawley, G., & Orfield, G. (2008). *The forgotten choice? Rethinking magnet schools in a changing landscape* [A Report to Magnet Schools of America]. The Civil Rights Project/Proyecto Derechos Civiles at UCLA.

Freeman v. Pitts, 503 U.S. 467 (1992).

Gamoran, A. (1996). Student achievement in public magnet, public comprehensive, and private city high schools. *Educational Evaluation and Policy Analysis, 18*(1), 1–18.

Gandara, P. & Hopkins, M. (Eds.). (2010). *Forbidden language: English learners and restrictive language policies.* Teachers College Press.

Gilliam, W. S. (2016). *Early childhood expulsions and suspensions undermine our nation's most promising agent of opportunity and social justice*. Robert Wood Johnson Foundation.

Gilliam, W. S., Maupin, A. N., Reyes, C. R., Accavitti, M., & Shic, F. (2016). *Do early educators' implicit biases regarding sex and race relate to behavior expectations and recommendations of preschool expulsions and suspensions?* Yale University Child Study Center.

Goldring, E. (2009). Perspectives on magnet schools. In M. Berends, M. Springer, D.

Ballou, & H. Walberg (Eds.), *Handbook of research on school choice* (pp. 361–378). Routledge.

Goldring, E. B. & Hausman, C. S. (1999). Reasons for parental choice of urban schools. *Journal of Education Policy, 14*(5), 469–490.

Goldring, E. & Smrekar, C. (2000). Magnet schools and the pursuit of racial balance. *Education and Urban Society*, *33*(1), 17–35.

Goldring, E. & Smrekar, C. (2002). Magnet schools: Reform and race in urban education. *The Clearing House*, *76*(1), 13–15.

Goldschmidt, P. & Martinez-Fernandez, J.-F. (2004). *The relationship between school quality and the probability of passing high-stakes performance assessments* (Technical Report No. 644). Center for Research on Evaluation Standards and Student Testing.

Gratto, S. D. (2002). Arts education in alternative school formats. *Arts Education Policy Review*, *103*(5), 17–24.

Green v. County School Board of New Kent County, 391 US 430 (1968).

Green, T. (2017). Community-based equity adults: A practical approach for educational leaders to support equitable community-school improvements. *Educational Administration Quarterly, 53*(1), 3-39.

Gregory, A., Nygreen, K., & Moran, D. (2006). The discipline gap and the normalization of failure. In P. A. Noguera & J. Y. Wing (Eds.), *Unfinished Business: Closing the Racial Achievement Gap in Our Schools* (pp. 121–150). Jossey-Bass, A Wiley Imprint.

Griffin, K. A., Allen, W. R., Kimura-Walsh, E., & Yamamura, E. K. (2007). Those who left, those who stayed: Exploring the educational opportunities of high-achieving black and Latina/o students at magnet and nonmagnet Los Angeles high schools (2001-2002). *Educational Studies*, *42*(3), 229–247.

Grooms, A. A. & Williams, S. M. (2015). The reversed role of magnets in St. Louis: Implications for Black student outcomes. *Urban Education*, *50*, 454–473.

Guiton, G. & Oakes, J. (1995). Opportunity to learn and conceptions of educational equality. *Educational Evaluation and Policy Analysis*, *17*(3), 323–336.

Guryan, J. (2004). Desegregation and black dropout rates. *The American Economic Review*, *94*(4), 919–943.

Haines, S. & Kilpatrick, C. (2007). Environmental education saves the day. *Science and Children*, *44*(8), 42–47.

Hannah-Jones, N. (2014, May 1). Lack of order: The Erosion of a once-great force for integration. *ProPublica*. http://www.propublica.org/article/lack-of-order-the-erosion-of-a-once-great-force-for-integration

Harris, D. N. (2006). *Ending the blame game on educational inequality: A study of "high-flying" schools and NCLB*. Great Lakes Center for Education Research and Practice.

Harris, J. C. (2015). *Modern-day magnet schools and modern-day needs: Are magnet schools improving educational equity for traditionally underserved students?* Michigan State University.

Hart v. Community School Board of Brooklyn, New York School District No. 21, 383 F. Supp. 699 (1974).

Hattie, J.A.C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to* *achievement.* Routledge.

Hausman, C. S. (2000). Principal role in magnet schools: Transformed or entrenched? *Journal of Education Administration*, *38*(1), 25–46.

Hausman, C. S. & Goldring, E. B. (2000a). Parent involvement, influence, and satisfaction in magnet schools: Do reasons for choice matter? *The Urban Review*, *32*(2), 105–121.

Hausman, C. S. & Goldring, E. B. (2000b). School community in different magnet program structures. *School Effectiveness and School Improvement*, *11*(1), 80–102.

Hausman, C. S. & Goldring, E. B. (2001). Teachers' ratings of effective principal leadership: A comparison of magnet and nonmagnet elementary schools. *Journal of School Leadership*, *11*(1), 399–423.

Hausman, C. S., Goldring, E. B., & Moirs, K. A. (1997, March 24). *Organizational capacity for school improvement: Teacher reports in magnet and nonmagnet schools*. Annual Meeting of the American Educational Research Association, Chicago, IL.

Haycock, K. (2001). Helping all students achieve: Closing the achievement gap. *Educational Leadership*, *58*(6), 6–11.

Haynes, N. M., Emmons, C., & Ben-Avie, M. (1997). School climate as a factor in student adjustment and achievement. *Journal of Educational and Psychological Consultation*, *8*(3), 321–329.

Helfrich, S.R. & Bosh, A.J. (2011). Teaching English language learners: Strategies for overcoming barriers. *The Educational Forum, 75*(3), 260-270.

Henig, J. R. (1990). Choice in public schools: An analysis of transfer requests among magnet schools. *Social Science Quarterly*, *71*(1), 69–82.

Herman, J. L. & Klein, D. C. D. (1997). *Assessing opportunity to learn: A California example* (CSE Technical Report No. 453). University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Herman, J. L., Klein, D. C. D., Wakai, S. T., & Heath, T. (1996). *Assessing equity in alternative assessment: An illustration of opportunity-to-learn issues* (CSE Technical Report No. 440). University of California, Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Hobson v. Hansen, 269 F. Supp. 401 (1967).

Holley-Walker, D. (2010). After unitary Status: Examining voluntary integration strategies for Southern school districts. *North Carolina Law Review*, *88*(3), 877–910.

Holme, J. J. (2002). Buying homes, buying schools: School choice and the social construction of school quality. *Harvard Education Review*, *72*(2), 177–205.

Howard, L.L. (2011). Does food insecurity at home affect non-cognitive performance at school?: A longitudinal analysis of elementary student classroom behavior. *Economics of Education Review*, *30*(1), 157-176.

Hurwitz, S., Perry, B., Cohen, E.D., & Skiba, R. (2019). Special education and individualized academic growth: A longitudinal assessment of outcomes for students with disabilities. *American Educational Research Journal, 57*(2), 576-611.

Jackson, L. D. (2007). A choice that works. *American School Board Journal*, *194*(12), 34–35.

Jacques, L. (1993, February 25). Proposed changes spark lively comments on magnet programs. *Miami Times*, 12A.

Jaffar, S. B. (2006). An alternative approach to measuring opportunity-to-learn in high school classes. *The Alberta Journal of Educational Research*, *52*(2), 107–126.

Jensen, E. (2009). *Teaching with poverty in mind: What being poor does to kids' brains and what schools can do about it.* ASCD.

Johnson, R. C. (2011). *Long-run impacts of school desegregation & school quality on adult attainments* (NBER Working Paper No. 16664). National Bureau of Economic Research.

Johnson, R. C. (2012). *The grandchildren of Brown: The long legacy of school desegregation* (p. Berkeley, CA) [Working Paper]. Goldman School of Public Policy, the University of California at Berkeley.

Kafer, K. (2012). *A Chronology of school choice in the U.S.* (pp. 1–18). Independence Institute.

Kahlenberg, R. D. (2001). *All together now: Creating middle-class schools through public school choice*. Brookings Institution Press.

Kahlenberg, R. D. (2011). Socioeconomic school integration: Preliminary lessons from more than 80 districts. In E. Frankenberg & E. DeBray (Eds.), *Integrating schools in a changing society: New policies and legal options for a multiracial generation* (pp. 167–186). University of North Carolina Press.

Kemple, J. J. & Scott-Clayton, J. (2004). *Career academies: Impacts on labor market outcomes and educational attainment*. Manpower Demonstration Research

Corporation.

Kemple, J. J. & Snipes, J. C. (2000). *Career academies: Impacts on students' engagement and performance in high school*. Manpower Demonstration Research Corporation.

Keyes v. School District No. 1, Denver, Colorado, 413 US 189 (1973).

Kini, T. & Podolsky, A. (2016). *Does teaching experience increase teacher effectiveness?: A review of the research.* Learning Policy Institute. https://learningpolicyinstitute.org/our-work/publications-resources/ does-teaching-experience-increase-teacher-effectiveness-review-research

Koedel, C., Betts, J. R., Rice, L. A., & Zau, A. C. (2009). The integrating and segregating effects of school choice. *Peabody Journal of Education*, *84*(2), 110–129.

Kurlaender, M. & Yun, J. T. (2001). Is diversity a compelling educational interest? Evidence from Louisville. In G. Orfield (Ed.), *Diversity challenged: Evidence on the impact of affirmative action*. Harvard Education Publishing Group.

Ladd, H.F. & Sorensen, L.C. (2015). *Do master's degrees matter?: Advanced degrees, career paths, and the effectiveness of teachers*. (Working Paper No. 136). National Center for Analysis of Longitudinal Data in Education Research.

Ladson-Billings, G. (2010). Foreword. In H. R. Milner, *Start where you are, but don't stay there: Understanding diversity, opportunity gaps, and teaching in today's classrooms* (pp. ix–xi). Harvard Education Press.

Larson, J. C. & Allen, B. A. (1988). *A microscope on magnet schools, 1983 to 1986, volume 2: Pupil and parent outcomes*. Montgomery County Public Schools Department of Educational Accountability.

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educational Researcher*, *31*(1), 3–12.

Lewis-Beck, C. & Lewis-Beck, M. (2016). *Applied regression: An introduction* (2nd ed.). SAGE Publications, Inc.

Linn, R. L. & Welner, K. G. (2007). *Race-conscious policies for assigning students to schools: Social science research and the Supreme Court cases*. National Academy of Education, Committee on Social Science Research Evidence on Racial Diversity in Schools.

Lutz, B. (2011). The end of court-ordered desegregation. *American Economic Journal: Economic Policy, 3*(2), 130–168.

Magnet Assistance Program enrollment and participation report. (1995). Charlotte-Mecklenburg Public Schools.

Magnet Schools Assistance Program, 34 C.F.R. § 280 (1985).

Magnet Schools of America. (2017). *A snapshot of magnet schools in America.* http://magnet.edu/files/Snapshot-of-Magnet-Schools-in-America.pdf

Manning, J. B. & Kovach, J. A. (2002). The continuing challenges of excellence and equity. In B. Williams (Ed.), *Closing the achievement gap: A vision for changing beliefs and practices* (2nd ed., pp. 25–47). Association for Supervision & Curriculum Development.

Martinez, V., Godwin, K., & Kemerer, F. R. (1996). Public school choice in San Antonio: Who chooses and with what effects? In B. Fuller, R. F. Elmore, & G. Orfield (Eds.), *Who Chooses, Who Loses?: Culture, Institutions, and the Unique Effects of School Choice* (pp. 50–69). Teachers College Press.

Martinez, V., Thomas, K., & Kemerer, F. R. (1994). Who chooses and why: A look at five choice plans. *Phi Delta Kappan*, *75*(9), 678–681.

Mattison, E. & Aber, M. S. (2007). Closing the achievement gap: The association of racial climate with achievement and behavioral outcomes. *American Journal of Community Psychology*, *40*, 1–12.

Mickelson, R. A. (2001). Subverting Swann: First- and second-generation segregation in the Charlotte-Mecklenburg Schools. *American Educational Research Journal*, *38*(2), 215–252.

Milliken v. Bradley, 418 US 717 (1974).

Milner, H. R. (2010). *Start where you are, but don't stay there: Understanding diversity, opportunity gaps, and teaching in today's classrooms*. Harvard Education Press.

Mintrop, H. & Zane, R. (2017). When the achievement gap becomes high stakes for special education teachers: Facing a dilemma with integrity. *Teachers College Record, 119*, 1-39.

Morales, H. (2010). Ignite zeal for STEM learning. *Techniques*, *85*(3), 22–24.

Morgan v. Kerrigan, 401 F. Supp. 216 (1976).

Morris, J. E. & Goldring, E. (1999). Are the magnet schools more equitable? An analysis of the disciplinary rates of African-American and White students in Cincinnati magnet and nonmagnet schools. *Equity & Excellence in Education*, *32*(3), 59–65.

Morrison, P. A. (1996). Forecasting enrollments during court-ordered desegregation. *Population Research and Policy Review*, *15*, 131–146.

Mujis, D. (2004). *Doing quantitative research in education with SPSS.* SAGE Publications, Inc.

National Center for Education Statistics. (2017). *English language learner (ELL) students enrolled in public elementary and secondary schools, by state: Selected years, fall 2000 through fall 2017*. https://nces.ed.gov/programs/digest/d19/tables/dt19_204.20.asp

NCES. (1991). *Education Counts: An indicator system to monitor the nation's educational health* [Report to the Special Study Panel on Education Indicators to the Acting Commissioner of Education Statistics]. National Center for Education Statistics.

Noguera, P. A. & Wing, J. Y. (2006). Unfinished business: Closing the achievement gap at Berkeley High School. In P. A. Noguera & J. Y. Wing (Eds.), *Unfinished Business: Closing the Racial Achievement Gap in Our Schools* (pp. 3–27). Jossey-Bass, A Wiley Imprint.

Oakes, J. (1986). Keeping track, part 1: The policy and practice of curriculum inequality. *Phi Delta Kappan*, *68*(1), 12–17.

Orfield, G. (2009). *Reviving the goal of an integrated society: A 21st century challenge*. The Civil Rights Project/Proyecto Derechos Civiles, UCLA.

Orfield, G., Frankenberg, E., Ee, J., & Kuscera, J. (2014). *Brown at 60: Great progress, a long retreat, and an uncertain future*. The Civil Rights Project/Proyecto Derechos Civiles at UCLA.

Orfield, G., Frankenberg, E., & Garces, L. M. (2008). Statement of American social scientists of research on school desegregation to the U.S. Supreme Court in Parents v. Seattle School District and Meredith v. Jefferson County. *Urban Review*, *40*, 96–136.

Orfield, G. & Lee, C. (2007). *Historic reversals, accelerating resegregation, and the need for new integration strategies*. The Civil Rights Project/Proyecto Derechos Civiles at UCLA.

Parents Involved in Community Schools v. Seattle School District No. 1, 551 US 701 (2007).

Penta, M. Q. (2001). *Comparing student performance at program magnet, year-round magnet, and non-magnet elementary schools* (Eye on Evaluation 1.27). Wake County Public School System, Evaluation and Research Department.

People Who Care v. Rockford Board of Education School District No. 205, 851 F. Supp. 905 (2001).

Petras, H., Masyn, K. E., Buckley, J. A., Ialongo, N. S., & Kellam, S. (2011). Who is most at risk for school removal? A multilevel discrete-time survival analysis of individual- and context-level influences. *Journal of Education Psychology, 103*(1), 223–237.

Pierce v. Society of Sisters, 268 U.S. 510 (1925).

Plessy v. Ferguson, 163 U.S. 537 (1896).

Poppell, J. B. & Hague, S. A. (2001). *Examining indicators to assess the overall effectiveness of magnet schools: A study of magnet schools in Jacksonville, Florida*. Annual Meeting of the American Educational Research Association, Seattle, WA.

Porter, A. (1995). The uses and misuses of opportunity-to-learn standards. *Educational Researcher, 24*(1), 21–27.

Reardon, S. F. (2016). School segregation and racial academic achievement gaps. *The*

Russell Sage Foundation Journal of the Social Sciences, *2*(5), 34–57.

Reardon, S. F., Grewal, E., Kalogrides, D., & Greenberg, E. (2012). Brown fades: The end of court-ordered school desegregation and the resegregation of American public schools. *Journal of Policy Analysis and Management*, *31*, 876–904.

Reardon, S. F. & Owens, A. (2014). 60 years after Brown: Trends and consequences of school segregation. *Annual Review of Sociology*, *40*, 199–218.

Reardon, S. F. & Rhodes, L. (2011). The effects of socioeconomic school integration plans on racial school desegregation. In E. Frankenberg & E. DeBray (Eds.), *Integrating schools in a changing society: New policies and legal options for a multiracial generation* (pp. 187–208). University of North Carolina Press.

Reardon, S. F. & Yun, J. T. (2003). Integrating neighborhoods, segregating schools: The retreat from school desegregation in the South, 1990–2000. *North Carolina Law Review*, *81*(4), 1563–1596.

Rhea, A. & Regan, R. (2007). *Magnet program review.* (Evaluation and Research Report 6.18). Wake County Public School System, Evaluation and Research Department.

Rickles, J., Ong, P. M., & Houston, D. (2004). *School integration and residential segregation in California: Challenges for racial equity*. The Ralph and Goldy Lewis Center for Regional Policy Studies at UCLA.

Robertson, K. (n.d.). *Increasing academic language knowledge for English language learner success.* ¡Colorín colorado! http://www.colorincolorado.org/article/increasing-academic-language-knowledge-english-language-learner-success

Ross v. Houston Independent School District, 699 F.2d 218 (1983).

Rossell, C. H. (2003). The desegregation efficiency of magnet schools. *Urban Affairs Review*, *38*(5), 697–725.

Rossell, C. H. (2005). Magnet schools: No longer famous but still intact. *Education Next*, *5*(2), 44–49.

Rothstein, R. (2004). *Class and schools: Using social, economic, and educational reform to close the black-white achievement gap*. Economic Policy Institute.

Rubin, B. C., Wing, J. Y., Noguera, P. A., Fuentes, E., Liou, D., Rodriguez, A. P., & McCready, L. T. (2006). Structuring inequality at Berkeley High. In P. A. Noguera & J. Y. Wing (Eds.), *Unfinished Business: Closing the Racial Achievement Gap in Our Schools* (pp. 29–86). Jossey-Bass, A Wiley Imprint.

Rubin, D. I. (2016). Growth in oral reading fluency of Spanish ELL students with learning disabilities. *Intervention in School and Clinic, 52*(1), 34-38

Saporito, S. (2003). Private choices, public consequences: Magnet school choice and segregation by race and poverty. *Social Problems*, *50*, 181–203.

Scherff, L. & Piazza, C. L. (2009). Why now, more than ever, we need to talk about opportunity to learn. *Journal of Adolescent & Adult Literacy, 52*(4), 343–352.

Schneider, M. & Buckley, J. (2002). What do parents want from school? Evidence from the Internet. *Educational Evaluation and Policy Analysis*, *24*(2), 133–144.

Schulte, A.C. & Stevens, J.J. (2015). Once, sometimes, or always in special education: Mathematics growth and achievement gaps. *Exceptional Children, 81*(3), 370-387.

Shonkoff, J. P. & Garner, A. S. (2012). The lifelong effects of early childhood adversity and toxic stress. *Pediatrics*, *129*(1), 232-46.

Siegel-Hawley, G. & Frankenberg, E. (2012). *Reviving magnet schools: Strengthening a successful choice option*. The Civil Rights Project/Proyecto Derechos Civiles, UCLA. http://www.escholarship.org/uc/item/5sv7r6cr

Siegel-Hawley, G. & Frankenberg, E. (2013). Designing choice: Structuring magnet schools for racial diversity. In G. Orfield & E. Frankenberg (Eds.), *Educational delusions? Why choice can deepen inequality and how to make it fair* (pp. 131–152). University of California Press.

Silver, D., Saunders, M., & Zarate, E. (2008). *What factors predict high school graduation in the Los Angeles Unified School District?* (California Dropout Research Project Report No. 14). University of California, Los Angeles.

Skiba, R. J., Horner, R. H., Chung, C.-G., Rausch, M. K., May, S. L., & Tobin, T. (2011). Race is not neutral: A national investigation of African American and Latino disproportionality in school discipline. *School Psychology Review*, *40*(1), 85–107.

Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *Urban Review*, *34*(4), 317–342.

Smith, D. (2002). Psychologist wins Nobel Prize. *Monitor on Psychology*, *33*(11), 22.

Smrekar, C. (2009). Beyond the tipping point: Issues of racial diversity in magnet schools following unitary status. *Peabody Journal of Education*, *84*(2), 209–226.

Smrekar, C. & Goldring, E. (1999). *School choice in urban America: Magnet schools and the pursuit of equity*. Teachers College Press.

Smrekar, C. & Goldring, E. (2000). *Social class isolation and racial diversity in magnet schools*. National Center for the Study of Privatization in Education, New York, NY.

Smrekar, C. & Goldring, E. (2011). Magnet schools, MSAP, and new opportunities to promote diversity. In E. Frankenberg & E. DeBray (Eds.), *Integrating schools in a changing society: New policies and legal options for a multiracial generation* (pp. 232–242). University of North Carolina Press.

Snyder, T. & Musu-Gillette, L. (2015, April 16). Free or reduced-price lunch: A proxy for poverty? *National Center for Education Statistics Blog*. https://nces.ed.gov/blogs/nces/post/free-or-reduced-price-lunch-a-proxy-for-poverty

Stanley, C. (1989). *Magnet schools: Fourteenth annual final report, 1968-89*. Houston Independent School District, Department of Research and Evaluation.

StataCorp (2019). *Stata 17*. StataCorp LLC. https://www.stata.com/why-use-stata/

Steele, L. & Eaton, M. (1996). *Reducing, eliminating, and preventing minority isolation in American schools: The impact of the Magnet Schools Assistance Program*. American Institutes for Research, U.S. Department of Education, Office of the Under Secretary.

Stein, S. J. (2000). Opportunity to learn as a policy outcome measure. *Studies in Educational Evaluation*, *26*, 289–314.

Stevens, F. I. (1993). *Opportunity to learn: Issues of equity for poor and minority students*. National Center for Education Statistics.

Stevens, F. I. (1996, April 8). *The need to expand the opportunity to learn framework: Should students, parents, and school resources be included?* Annual Meeting of the American Educational Research Association, New York, NY.

Swann v. Charlotte-Mecklenburg Board of Education, 402 US 1 (1971).

Swanson, C. B. (2004). *Who graduates? Who doesn't? A statistical portrait of public high school graduation, class of 2001*. Education Policy Center, The Urban Institute.

Teske, P., Fitzpatrick, J., & Kaplan, G. (2007). *Opening doors: How low-income parents search for the right school*. University of Washington, Daniel J. Evans School of Public Affairs, Center on Reinventing Public Education.

Texas Education Agency. (2017, February). *District type glossary of terms, 2015-2016.* https://tea.texas.gov/reports-and-data/school-data/district-type-data-search/district-type-glossary-of-terms-2015-16

Texas Education Agency. (2019). *2018-2019 Texas academic performance reports*. https://rptsvr1.tea.texas.gov/perfreport/tapr/2019/index.html

Texas Education Agency. (2020a). *San Antonio ISD: Finance summary 2018-2019.* https://txschools.gov/districts/015907/finance

Texas Education Agency. (2020b). *Socorro ISD: Finance summary 2018-2019.* https://txschools.gov/districts/071909/finance

Texas Education Agency, Division of Research and Analysis, Office of Governance and Accountability. (2020, August). *Enrollment in Texas public schools, 2019-20.* https://tea.texas.gov/sites/default/files/enroll_2019-20.pdf

Texas Education Agency. (n.d.a). *All about the STAAR test.* https://www.texasassessment.com/staar/families/all-about-the-staar-test/#:~:text=STAAR%20is%20the%20state's%20testing,are%20given%20throughout%20the%20year.

Texas Education Agency. (n.d.b). *Comprehensive glossary: 2018-19 Texas academic performance report*. https://rptsvr1.tea.texas.gov/perfreport/tapr/2019/glossary.pdf

Texas Education Agency. (n.d.). *Reports and data*. https://tea.texas.gov/reports-and-data

Texas Education Agency, Division of Research and Analysis, Office of Governance and Accountability. (2020, August). *Enrollment in Texas public schools, 2019-20.* https://tea.texas.gov/sites/default/files/enroll_2019-20.pdf

The Civil Rights Act of 1964, 88–352 Pub.L. (1964).

Thurlow, M.L., Wu, Y., Lazarus, S.S,. & Ysseldyke, J.E. (2016). Special education-non-special education achievement gap in math: Effects of reporting methods, analytical techniques, and reclassification. *Exceptionality, 24*(1), 32-44.

Tuttle v. Arlington County School Board, 195 F. 3d 698 (1999).

United States Census Bureau. (2019, December 30). *State population totals and components of change: 2010-2019*. https://www.census.gov/data/tables/time-series/demo/popest/2010s-state-total.html#par_textimage_1574439295

United States Department of Agriculture. (2018, May 8). *Child nutrition programs: Income eligibility guidelines.* https://www.federalregister.gov/documents/2018/05/08/2018-09679/child-nutrition-programs-income-eligibility-guidelines

U.S. Department of Education. (2008). New Haven voluntary public school choice: Expanding options, improving achievement. *The Education Innovator, 6*(3), 1–3.

U.S. Department of Education. (2017). *Magnet Schools Assistance program description*. Magnet Schools Assistance. https://www2.ed.gov/programs/magnet/index.html

U.S. Department of Education, Office of Civil Rights. (2014). *Data snapshot: Early childhood education* (Issue brief #2; Civil Rights Data Collection).

Walsh, M. (2007). Use of race a concern for magnet schools; decision in student-assignment cases came as U.S. reviewed grant applications. *Education Week*, *27*(10), 8.

Walton, M. & Ford, E. (2014). *Magnet Schools Assistance Program grantee data analysis report: Minority group isolation and student academic achievement*. Magnet Schools Assistance Program Technical Assistance Center. https://www2.ed.gov/programs/magnet/granteedata.pdf

Wang, J. (1998). Opportunity to learn: The impacts and policy implications. *Educational Evaluation and Policy Analysis*, *20*(3).

Wang, J. & Herman, J. L. (2017). Magnet schools: History, description, and effects. In R. Fox & N. Buchanan (Eds.), *The Wiley Handbook of School Choice* (pp. 158–179). John Wiley and Sons.

Wang, J., Schweig, J. D., & Herman, J. L. (2014). *Is there a magnet school effect? Using meta-analysis to explore variation in magnet school success* (CRESST Report No. 843). National Center for Research on Evaluation, Standards, & Student Testing at the University of California, Los Angeles.

Wang, J., Schweig, J. D., & Herman, J. L. (2017). Is there a magnet-school effect? A multisite study of MSAP-funded magnet schools. *Journal of Education for Students Placed at Risk*, *22*(2), 77–99.

Wehlage, G. G. & Smith, G. A. (1992). Building new programs for students at risk. In F. M. Newmann (Ed.), *Student engagement and achievement in American secondary* *schools* (pp. 92–118). Teachers College Press.

Wei, X., Lenz, K.B., & Blackorby, J. (2012). Math growth trajectories of students with disabilities: Disability category, gender, racial, and socioeconomic status differences from ages 7 to 17. *Remedial and Special Education, 34*(3), 154-165.

Weiss, C. H. (1996). Foreword. In B. Fuller, R. F. Elmore, & G. Orfield (Eds.), *Who Chooses, Who Loses?: Culture, Institutions, and the Unique Effects of School Choice* (pp. vii–viii). Teachers College Press.

Welner, K. G. (2006). K-12 race-conscious student assignment policies: Law, social science, and diversity. *Review of Educational Research*, *76*(3), 349–382.

Wessmann v. Gittens, 160 F. 3d 790 (1998).

West, K. C. (1994). A desegregation tool that backfired: Magnet schools and classroom segregation. *Yale Law Journal*, *103*(8), 2567–2592.

Wiley, D. E. & Yoon, B. (1995). Teacher reports on opportunity to learn: Analyses of the 1993 California Learning Assessment System (CLAS). *Educational Evaluation and Policy Analysis*, *17*(3), 355–370.

Williams, B. (2002a). Introduction. In B. Williams (Ed.), *Closing the achievement gap: A vision for changing beliefs and practices* (2nd ed., pp. 1–12). Association for Supervision & Curriculum Development.

Williams, B. (2002b). What else do we need to know and do? In B. Williams (Ed.), *Closing the achievement gap: A vision for changing beliefs and practices* (2nd ed., pp. 13–24). Association for Supervision & Curriculum Development.

Wing, J. Y. (2006). Integration across campus, segregation across classrooms: A close-up look at privilege. In P. A. Noguera & J. Y. Wing (Eds.), *Unfinished Business:*

Closing the Racial Achievement Gap in Our Schools (pp. 87–119). Jossey-Bass, A Wiley Imprint.

Witte, J. F. & Walsh, D. T. (1990). A systematic test of the effective schools model. *Educational Evaluation and Policy Analysis*, *12*(2), 188–212.

Wu, S.-C., Pink, W., Crain, R., & Moles, O. (1982). Student suspension: A critical reappraisal. *Urban Review*, *14*(4), 245–303.

Yang, N. Y., Li, Y. H., Tompkins, L. J., & Modarresi, S. (2005). *Using the multiple-matched-sample and statistical controls to examine the effects of magnet school programs on the reading and mathematics performance of students*. Annual Meeting of the American Educational Research Association, Montreal, Canada.

Yu, C. M. & Taylor, W. L. (1997). *Difficult choices: Do magnet schools serve children in need?* Citizens' Commission on Civil Rights.

Appendix A

Dependent Variable Descriptive Statistics

	Observations	Mean	Standard Deviation	Minimum	Maximum
2019 Fifth Grade Reading STAAR Test					
White/African American Gaps					
Approaches	194	.074	.145	-0.56	0.50
Meets	194	.184	.221	-0.50	0.90
Masters	194	.190	.226	-0.63	0.75
White/Hispanic Gaps					
Approaches	300	.041	.110	-0.41	0.32
Meets	300	.134	.189	-0.41	0.73
Masters	300	.144	.186	-0.36	0.72
2019 Fifth Grade Math STAAR Test					
White/African American Gaps					
Approaches	194	.084	.134	-0.25	0.77
Meets	194	.184	.216	-0.44	0.73
Masters	194	.211	.223	-0.44	0.80
White/Hispanic Gaps					
Approaches	300	.030	.102	-0.37	0.38
Meets	300	.100	.168	-0.42	0.60
Masters	300	.125	.186	-0.40	0.63
2019 English End-of-Course STAAR Tests					
White/African American Gaps					
Approaches	120	.117	.155	-0.27	0.62
Meets	120	.170	164	-0.32	0.63

Masters	120	.085	.111	-0.14	0.45
White/Hispanic Gaps					
Approaches	134	.075	.121	-0.24	0.51
Meets	134	.135	.125	-0.18	0.55
Masters	134	.081	.104	-0.13	0.39
2019 Algebra I End-of-Course STAAR Test					
White/African American Gaps					
Approaches	138	.019	.112	-0.36	0.33
Meets	138	.053	.141	-0.50	0.41
Masters	138	.060	.164	-0.50	0.66
White/Hispanic Gaps					
Approaches	194	-.002	.083	-0.38	0.34
Meets	194	.020	.106	-0.36	0.34
Masters	194	.038	.137	-0.63	0.43

Appendix B

Significant Control Variable Coefficients

Variable	Model	Coefficient	*t*	*P* > |*t*|	R^2
Student Demographic Control Variables					
Economic disadvantage percentage	English EOC, masters, White/African American	-0.2100	-3.12	0.002	0.644
Economic disadvantage percentage	English EOC, masters, White/Hispanic	-0.1713	-2.39	0.019	0.486
English learner percentage	5th math, masters, White/Hispanic	-0.2644	-2.63	0.009	0.193
English learner percentage	English EOC, meets, White/African American	-0.4442	-2.14	0.035	0.458
Special education percentage	5th reading, meets, White/Hispanic	0.9277	2.15	0.032	0.213
Special education percentage	English EOC, approaches, White/Hispanic	0.9020	2.13	0.036	0.290
Special education percentage	English EOC, meets, White/Hispanic	0.9350	2.18	0.031	0.320
Enrollment	5th reading, meets, White/Hispanic	-0.0002	-2.96	0.003	0.213
Enrollment	English EOC, meets, White/African American	0.00005	2.19	0.031	0.458
Enrollment	Algebra EOC, meets, White/African American	0.00005	3.13	0.002	0.440
Educational Input Control Variables					
Average ELA class size	English EOC, approaches, White/African American	-0.0116	-2.35	0.021	0.350

Average ELA class size	English EOC, meets, White/ African American	-0.0111	-2.32	0.023	0.458
School District Control Variables					
Austin ISD	5th reading, meets, White/Hispanic	0.1178	2.52	0.012	0.213
Austin ISD	5th reading, masters, White/Hispanic	0.1091	2.35	0.020	0.197
Austin ISD	5th math, approaches, White/African American	0.1354	2.85	0.005	0.213
Austin ISD	5th math, meets, White/African American	0.1763	2.26	0.025	0.187
Austin ISD	5th math, masters, White/African American	0.3066	3.91	0.000	0.231
Austin ISD	5th math, meets, White/Hispanic	0.0993	2.33	0.021	0.174
Austin ISD	5th math, masters, White/Hispanic	0.1412	3.03	0.003	0.193
Austin ISD	Algebra EOC, meets, White/African American	0.1570	2.68	0.008	0.440
Houston ISD	5th reading, meets, White/Hispanic	0.1357	2.69	0.008	0.213
Houston ISD	5th math, masters, White/Hispanic	0.1186	2.36	0.019	0.193
San Antonio ISD	5th math, approaches, White/Hispanic	0.1646	2.09	0.037	0.106
Socorro ISD	5th math, approaches, White/African American	-0.1285	-2.28	0.024	0.213
Socorro ISD	5th math, masters, White/African American	0.1850	1.99	0.048	0.231
Socorro ISD	Algebra EOC, masters, White/African American	-0.2440	-2.59	0.011	0.307

Ysleta ISD	5th math, meets, White/African American	-0.3918	-2.46	0.015	0.187
Ysleta ISD	English EOC, meets, White/African American	0.2389	2.20	0.031	0.458

Note. All findings included in this table occurred in the context of the regressions conducted in the context of the second research question.