DOES THE PROPORTION OF WHITE STUDENTS PREDICT DISCIPLINE DISPARITIES? A NATIONAL, SCHOOL-LEVEL ANALYSIS OF SIX RACIAL/ETHNIC STUDENT GROUPS

by

Nicholas Paul Triplett

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in Curriculum and Instruction

Charlotte

2018

Approved by:

_____

Dr. Bettie Ray Butler

_____

Dr. Chance Lewis

_____

Dr. Tina Heafner

_____

Dr. Claudia Flowers

ABSTRACT

NICHOLAS PAUL TRIPLETT. Does the proportion of white students predict discipline disparities? A national, school-level analysis of six racial/ethnic student groups (Under the direction of DR. BETTIE RAY BUTLER)

Racial/ethnic disparities in school discipline have existed for decades (Children's Defense Fund, 1975), and have had a disproportionately negative effect on the educational attainment, social/emotional well being and life chances of students of color. While past research has found that school racial/ethnic balance is amongst the most powerful predictors of the student-level risk and severity of school discipline (Rocque & Paternoster, 2011; Skiba, Chung, Trachok, Baker, Sheya & Hughes, 2014; Welch & Payne, 2010), only a handful of studies have examined how school racial/ethnic balance affects discipline gaps between students of color and their White counterparts (Freeman & Steidl, 2016; Thornton & Trent, 1988). However, careful analysis of previous scholarship suggests that students of color attending schools with higher proportions of White students may be particularly vulnerable to racialized and overly punitive discipline practices (Eitle & Eitle, 2004; Welch & Payne, 2010). The primary purpose of this study was to examine the relationship between school-level racial/ethnic disparities in suspension and the proportion of White students in a school. Regression analysis was used to analyze a nationally representative sample of K-12 schools for each U.S. Census designated student groups of color: American Indian/Alaska Native, Asian, Black, Latinx, Native Hawaiian/Pacific Islander, Two or more Races. Analysis was guided by critical race and intergroup threat theories, which explored the ways that structural racism and perceptions of intergroup threat can produce ideologies, dispositions and actions that

deny historically marginalized youth the full benefit of public education (Ladson-Billings & Tate, 1995; Welch & Payne, 2010). Descriptive findings demonstrated that the problem of inequitable suspension rates was perhaps more acute and widespread than previously appreciated. This study provided some of the first empirical evidence of elevated risk of suspension for Asians, Native Hawaiian/Pacific Islanders, students of two or more Races, and students of color in predominantly-White schools. Inferential analysis offered support for two important emerging findings in the school discipline literature. First, analysis indicated that elementary schools tend to have consistently larger racial/ethnic discipline gaps than middle and high schools. Secondly, multivariate analysis showed that the best predictors of school level racial/ethnic discipline gaps were different from those that have best predicted student-level incidence and severity of discipline in previous scholarship. Collectively, results contributed empirical evidence that can help identify the school contexts in which students of color appear to be at elevated risk of suspension, what kinds of reforms might produce more equitable discipline outcomes, where such reforms should be implemented, and how stakeholders can mitigate the irreparable harm caused by racialized discipline policies and practices. Directions for future research are provided, along with recommendations for policy and practice.

ACKNOWLEDGEMENTS

I am incredibly grateful to my wife and children, who have made innumerable sacrifices in support of my doctoral journey. It is only through their love and understanding that I have been able to be a husband, father, and doctoral student simultaneiously. I must thank Kate Brown and Cara Zupancic for always being willing to sacrifice thier time to make sure my children made it to school and dinner was ready on time. I am deeply indebted to my committee members for the time and expertise they provide to make this project successful. Dr. Chance Lewis, thank you for seeing potential in me in our very first class together; for standing behind me as I learned to stand up; for keeping my work focused on what is important; and for your ability to distill the "nuggets" from the noise. Dr. Tina Heafner, thank you for accepting and treating me as a partner in your work; for providing me with so many professional opportunities; for showing me that you can have numerous important projects going on at one time, and still do a fine job on all of them; and for modelling the ways that highly skilled, highly productive working scholars get the job done. Dr. Claudia Flowers, thank you for graciously stepping into this project mid-stream; for providing me with sound and timely methodological advice at several critical, and highly stressful points; for applying a light touch to my statistical work and allowing me to build the study that I envisioned; and for recognizing that this work will continue beyond this dissertation. Dr. Bettie Ray Butler, thank you for accepting me as your first dissertation mentee; for teaching me so much about our shared interests of race/ethnicity and school discipline; for helping me find my voice and letting me add my voice to yours on these important matters; and for always working behind the scenes to promote my best interests, even if I didn't always realize it.

TABLE OF CONTENTS

LIST OF TABLES.....xii

LIST OF FIGURES.....xiii

CHAPTER 1: INTRODUCTION.....1

Statement of the Problem.....3

Purpose of the Study.....6

Theoretical Framework.....6

Critical Race Theory.....7

Intergroup Threat Theory.....8

Overview of Context and Methods.....10

Research Question.....12

Significance Statement.....13

Summary…….....14

Definition of Terms.....15

Assumptions.....16

CHAPTER 2: REVIEW OF LITERATURE.....17

Overview….....17

Discipline Disparities.....18

Who Experiences Disproportional Discipline?.....18

The Consequences of Disproportional Discipline.....18

Academic Achievement & Attainment.....19

Social & Emotional Factors.....19

Juvenile Justice & the School-to-Prison Pipeline.....20

Race, Ethnicity & School Discipline....21
Student Variables....22
Socioeconomic Status....23
Behavioral Differences....24
School and School Authorities Variables....26
Zero Tolerance Policies....27
The Attitudes & Dispositions of School Authorities....30
Demographic Composition....34
Zero Tolerance Policies....31
Discipline in Predominantly-White Schools....35
Racial Balance....36
Teacher Characteristics....38
School Wealth & Resources....40
Reform and Intervention....43
Theoretical Framework....45
Critical Race Theory....45
Intergroup Threat Theories....47
Summary & Conclusion....49
CHAPTER 3: METHOD....51
Research Question....52
Hypotheses....52
Data....53
Variables....53

Dependent Variable.....53
Focal Independent Variable.....55
Control Variables.....56
Proportion of Female Students.....56
Proportion of Students Identified as Disabled.....56
Proportion of Students Identified as Limited English Proficient...57
Proportion of Novice Teachers.....58
Full-Time Counselors.....59
Proportion of Students Eligible for Free/Reduced Lunch.....59
School Level.....60
School Locale.....61
Model Fitting.....62
Meaningful and Stable Estimates of the Dependent Variable.....62
Normality of the Dependent Variable.....63
Sample. & Selection.....64
Criterion 1.....64
Criterion 2.....65
Criterion 3.....66
Criterion 4.....66
Data Screening.....68
Analysis.....70
Assumptions.....71
Sample Size.....71

Normality of the Dependent Variable.....71
Normal Distribution of Residual Errors.....72
Univariate Outliers.....72
Multivariate Outliers.....73
Multicollinearity.....74
Linearity.....74
Homoscedasticity.....74
Independence of Observations/Residuals.....74
CHAPTER 4: RESULTS.....76
Descriptive Findings.....76
American Indian/Alaska Native.....77
Asian.....79
Black.....79
Latinx.....81
Native Hawaiian/Pacific Islander.....84
Two or More Races.....84
Between & Among Group Comparisons.....87
Regression.....91
American Indian/Alaska Native…….....92
Asian…….....93
Black…….....94
Latinx…….....95
Native Hawaiian/Pacific Islander…….....96

Two or More Races…….....97

Between & Among Group Comparisons.....97

Proportion White.....99

Proportion Female.....99

Proportion Disability.....99

Proportion LEP.....99

Proportion Novice Teachers.....100

Full Time Counselors.....100

Proportion FRL.....100

School Level.....101

School Locale.....101

Model Fit (R^2) 102

Summary of Notable Findings.....103

CHAPTER 5: DISCUSSION & IMPLICATIONS.....106

Discussion of Descriptive Findings.....109

Racial/Ethnic Discipline Disparities.....109

The Distribution of Racial/Ethnic Disparities in Suspension.....112

Discipline in Predominantly-White Schools.....114

Discussion of Inferential Findings.....116

Proportion of White Students.....116

Control Variables.....118

Proportion of Students Eligible for Free/Reduced Lunch.....118

School Level.....120

Gender Balance....122

Disability Status....122

Limited English Proficiency....123

Novice Teachers....123

Full Time Counselors....124

School Locale....124

Delimitations....125

Limitations....127

Recommendations for Policy & Reform....127

Federal Discipline Accountability....129

Separating Subjective and Objective Disciplinary Processes....133

Slow Discipline & Restorative Practices....135

Future Research....137

Summary and Conclusions....140

REFERENCES....143

APPENDIX A: Sample Details by Racial/Ethnic Group....161

APPENDIX B: Sample Size by Predominantly-White Status……....162

APPENDIX C: School Level Relative Risk Ratio of Suspension by the Proportion of White Students (Quartiles)....163

APPENDIX D: The Proportion 0f White Students by Racial/Ethnic Group Sample....164

LIST OF TABLES

Table 1: Excluded cases by reason and final sample.....68

Table 2: American Indian/Alaska Native means, standard deviations, and intercorrelations among all variables (N=4017).....78

Table 3: Asian means, standard deviations, and intercorrelations among all variables (N= 7786) 80

Table 4: Black means, standard deviations, and intercorrelations among all variables (N=29,196) 82

Table 5: Latinx means, standard deviations, and intercorrelations among all variables (N=29,526) 83

Table 6: Native Hawaiian/Pacific Islander means, standard deviations, and intercorrelations among all variables (N=1089) 85

Table 7: Two or More Races means, standard deviations, and intercorrelations among all variables (N=17853*)*.....86

Table 8: Mean values for variables by racial/ethnic group.....88

Table 9: School level relative risk ratio of suspension by predominantly-White status.....89

Table 10: Percentage of schools with racial/ethnic (dis)proportionality and relative risk ratios…….....90

Table 11: The impact of the proportion of white students in a school American Indian/Alaska Native risk of suspension.....93

Table 12: The impact of the proportion of white students in a school on Asian risk of suspension.....94

Table 13: The impact of the proportion of white students in a school on Black risk of suspension.....95

Table 14: The impact of the proportion of white students in a school on Latinx risk of suspension.....96

Table 15: The impact of the proportion of white students in a school on Native Hawaiian/Pacific Islander risk of suspension.....96

Table 16: The impact of the proportion of white students in a school on students

of two or more races risk of suspension.....97

Table 17: Regression outcomes across racial/ethnic groups.....98

LIST OF FIGURES

Figure 1: Relative Risk Ratio (RRR) Calculation for Black Students.....54

CHAPTER 1: INTRODUCTION

School discipline has emerged as a critical arena in the quest for racial/ethnic equity in American schools. Concern over racial/ethnic disparities in school discipline has reached the highest levels of educational policymaking as evidenced by the 2014 joint release of discipline reform guidelines by the U.S. Departments of Justice and Education, which pointed to evidence that students of color tend to be disciplined more frequently and more harshly than their White counterparts (U. S Department of Justice & U.S. Department of Education, 2014). While these federal school discipline guidelines are voluntary, school districts and state departments of education are increasingly acting to address discipline inequity (Minneapolis Public Schools, 2014; Oregon Department of Public Education, 2014; U.S. Department of Education, 2016a). In addition, federal education legislation recently recognized disparities in school discipline and the overuse of punitive punishment as important factors in educational achievement and attainment (Every Student Succeeds Act, 2015; Triplett, Bryant, Brown, Steele, Ardrey, Allen, & Lewis, 2016). These actions have brought much needed attention to equity in school discipline, however they are merely the most recent iteration of similar calls for reform in school discipline that stretch back over four decades (Children's Defense Fund, 1975). Despite increased attention, it remains unclear what kinds of reforms can most effectively reduce discipline disparities, and where to target such reforms (Skiba, Arredondo & Rausch, 2014).

This study proceeded in recognition of several research findings, which taken together, signal the need for greater attention to the racial/ethnic inequity in the application of exclusionary school discipline. First, studies of school discipline have been remarkably unanimous and consistent in concluding that students of color are over-selected for punitive school discipline, even after accounting for variations in student socioeconomic class and misbehavior rates (Children's Defense Fund, 1975; Fabelo, Thompson, Plotkin, Carmichael, Marchbanks & Booth, 2011; Wallace, Goodkind, Wallace & Bachman, 2008). The over-selection of students of color for discipline has been shown to diminish educational attainment, harm social/emotional well being and increase interactions with criminal and juvenile justice systems (Fabelo et al., 2011; Losen & Martinez, 2013; Wallace et al., 2008). Disproportional and racialized school discipline have become imbued with a universal, "normalized" (Watts & Erevelles, 2004, p. 293) quality that positions discipline disparities alongside the numerous structural forms of racial/ethnic discrimination operating every day in American social, legal and educational institutions (Bonilla-Silva, 2010; Ladson-Billings & Tate, 1995). Second, previous studies have also recognized that diverse, metropolitan school contexts, schools and districts that currently serve, or have historically served, predominantly-White student populations exhibit a collection of traits that may make students of color particularly vulnerable to discriminatory discipline practices (Betts, Reuban & Dananberg, 2000; Eitle & Eitle, 2004; Raffaele Mendez, Knoff & Ferron, 2002). Third, demographic trends indicate that growing numbers of students of color will be attending U.S. schools in general, and that these students of color are likely to constitute a larger proportion of students in what are currently predominantly-White institutions in the

future (Orfield & Frankenberg, 2014; U.S. Census Bureau, 2015). Taken together, these previous findings suggest that, in the absence of significant intervention, the harm done by inequitable and racialized school discipline practices will continue to disproportionately affect students of color in American schools.

The present study was guided by two theories that framed differential outcomes related to both the structural (educational institutions) and the individual (classrooms, school authorities) factors associated with racial/ethnic discipline disproportionality: critical race theory (Bell, 1973) and intergroup threat theory (Blalock, 1967; Blumer, 1958; Stephan & Stephan, 1996). This framework emphasized how discipline outcomes in predominantly-White schools can serve as a site for a) the application of negative racial/ethnic biases against students of color and b) the use of social control measures (school discipline) to maintain the institutional, social and cultural dominance of Whites in schools. Using school level demographic and discipline data from almost K-12 public schools in the U.S., this study sought to determine how the proportion of White students influenced the discipline gap between students of color and their White counterparts in the same schools. Findings provided evidence that can help educational stakeholders identify the traits of schools in which students of color are at an elevated risk of being over-selected for suspension. As a result of these findings, this study proposed policies, practices and interventions intended to reduce the negative consequences associated with the inequitable application of suspension.

Statement of the Problem

Since the Children's Defense Fund (1975) documented racial/ethnic disparities for African American students, numerous studies have remained remarkably consistent in

their finding that African American students are over-selected for punitive and exclusionary discipline across a variety of school contexts (Krezmien, Leone & Achilles, 2006; Raffaele Mendez & Knoff, 2003; Skiba et al., 2014; Wallace et al., 2008; Welch & Payne, 2010). Indeed, just as the Children's Defense Fund (1975) used national data from the Office of Civil Rights (OCR) to conclude that African American high school students were three times as likely as White students to be suspended, analysis of data from the same source revealed that African Americans remained between three to six times more likely (depending on school district) to receive suspension over 30 years later (Public Counsel Law Center, 2012). While the degree of racial/ethnic disparity and the consistency of the empirical evidence are highest amongst African Americans, exclusionary discipline is also disproportionately applied to American Indians (Brown & DiTillio, 2013; Krezmien et al., 2006; Losen & Gillespie, 2012; Rocque, 2010) and Latinx populations (Arcia, 2007; Peguero & Shekarkhar, 2011; Skiba, Horner, Rausch, & Tobin, 2011). The overuse and inequitable application of exclusionary discipline has damaging effects on the academic outcomes, social/emotional well being and ultimately the life chances of students that are most often the target of such punishments (Fabelo et al., 2011; Losen & Martinez, 2013; Wallace et al., 2008).

Some studies have found that the racial/ethnic composition of schools was a powerful predictor of increased incidence and punitiveness of discipline in schools (Skiba et al., 2014; Rocha & Hawes, 2009; Welch & Payne, 2010). A concurrent line of research has examined discipline *rates* between racial/ethnic groups, and has suggested that disparities between Black and White suspension may be greater in more highly resourced districts that tend to exhibit higher proportions of White students (Eitle & Eitle, 2004;

Rausch & Skiba, 2006; Wallace et al., 2008). However, while previous research has made it plain that students of color experience the highest incidence of suspension in diverse, urban schools, few studies have directly examined school discipline in more homogenous, historically White school contexts (Monroe, 2013); and even fewer studies have modeled the discipline gap to determine what school traits predict a higher risk of suspension for students of color as compared to White students (Eitle & Eitle, 2004; Freeman& Steidl, 2016; Thornton & Trent, 1988).

Meanwhile, demographic trends have shown that increasing numbers of students of color will be attending school in these more homogenous, historically White contexts (Orfield & Frankenberg, 2014; U.S. Census Bureau, 2015). Indeed, data used in this study (from the U.S. Department of Education Civil Rights Data Collection) indicated that approximately 1 in 4 public school students of color (about 6.5 million) attended a predominantly-White school in 2013-2014. Despite these demographic realities, little is known about how schools of different levels (elementary vs secondary, schools situated in different locales (city, suburban, town, rural), and schools with different demographic compositions apply suspension to various racial/ethnic groups. Empirical analysis of these intricacies has the potential to help stakeholders predict and track where students of color may be at heightened risk of experiencing not only an increased incidence of discipline, but also more racialized and punitive forms of punishment. This, in turn, can help clarify where intervention might be extended, and how stakeholders might produce school discipline policies and practices that are less harmful to students of color.

Purpose of the Study

The purpose of this study was to determine how the proportion of White students in a school is related to the degree of racial/ethnic disproportionality in suspension, which can increase empirical understanding of what kinds of schools tend to have the widest gaps in discipline between students of color and their White counterparts. Findings were also intended to provide analysis of school-level discipline differences in large, national samples, across all U.S Census designated racial/ethnic groups, including several previously under-researched student populations (i.e. students of color predominantly-White schools, Native Hawaiian/Pacific Islanders, students of two or more races.

Theoretical Framework

This study uses two related theoretical traditions to explore the ways that race and ethnicity frame our understanding of outcomes in school discipline. Critical race theory (Bell, 1973) positions schools as an important structural context where students are rewarded for displaying White cultural norms and sanctioned for cultural traits (e.g., dress, speech patterns, patterns of social interaction) associated with other cultural groups (Gay, 2006; Ladson-Billings, 1998; Neal, McCray, Webb-Johnson & Bridgest, 2003; Townsend, 2000). Intergroup threat theory (ITT) (Blalock, 1967; Blumer, 1958; Stephan & Stephan, 1996) explores how the presence of non-dominant social groups can be perceived as a criminal, social or cultural threat by dominant groups, resulting in prejudice, negative stereotypes and increased application social control measures as a means of maintaining social dominance.

While most disciplinary actions begin in the classroom between individuals or groups of individuals (Skiba, Michael, Nardo, & Peterson, 2002), differential discipline

outcomes also reflect the racial/ethnic attitudes and dispositions of educational institutions and prevailing social structures (Skiba, 2000; Watts & Erevelles, 2004; Wu, Pink, Crain, & Moles, 1982). Studies examining the effect of race in schooling have found evidence of racial threat, stereotyping and cultural bias in school discipline, teacher expectations and student achievement outcomes (Rocque & Paternoster, 2011; Van den Bergh, Denessen, Hornstra, Voeten & Holland, 2010; Welch & Payne, 2010). The use of both CRT and ITT allows this study to address the role of race in the administration of school discipline by delineating both the macro (structural) and micro (individual, classroom) aspects of inequitable disciplinary outcomes.

Critical Race Theory

Critical race theory emerged in American legal scholarship in the late 1980s (Bell, 1995) and has spread to other disciplines, including education (Ladson-Billings & Tate, 1995; Solarzano & Bernal, 2001). CRT positions racism as an endemic, structural force rooted in a history of violence and oppression, and built into ideological and social institutions (Bell, 1973; Ladson-Billings & Tate, 1995; Watts & Erevelles, 2004). While purposive individual acts of discrimination are an important component of racism, CRT focuses on how fundamental economic, legal and educational institutions function to preserve White control over power and material resources in society (Parker & Lynn, 2002; Sleeter, 2012). Scholars of critical race theory reject the purported race neutrality or colorblindness of prevailing school discipline policies and practices and foregrounds how structural and institutional factors can contain endemic discriminatory elements that serve to punish students of color based on racialized interpretations of student behavior (Ladson-Billings & Tate, 1995; Rocque & Paternoster, 2011; Payne & Welch, 2010).

Critical race analysis also challenges the ability of school authorities to practice school discipline in a race-neutral manner based on the operation of racial/ethnic deficit ideologies, cultural mismatch and widely held associations between Black males and predatory criminality (Giroux, 2003, Ladson-Billings & Tate, 1995, Payne & Welch, 2010). School discipline represents an arena in which socially constructed definitions of students of color as violent or deviant can "leak" into discipline outcomes to justify punishment and exclusion (Butler, Robinson & Walton, 2014; Gregory & Weinstein, 2008; Watts & Erevelles, 2004). This study uses CRT to frame how structural factors, such as discipline policies, social/economic arrangements, residential patterns, and laws and legislation help to explain pervasive and persistent racial/ethnic disparities in school discipline.

Intergroup Threat Theory

Early iterations of intergroup threat theory (ITT) hypothesized that increasing numbers of individuals from non-dominant social groups within a bounded system (neighborhood, school, etc.) are perceived to present an economic, political or criminal threat to dominant social groups (Bobo, 1988; Sherif & Sherif, 1969). The perception of such *realistic threats* (or material threats) is theorized to produce negative attitudes toward non-dominant groups and leads to the imposition of punitive social controls in order to maintain social dominance. A prominent example of this tradition is racial threat theory (RTT) (Blalock, 1967), which emerged from American legal studies based on research findings that showed increasing concentrations of African Americans within a neighborhood were consistently and positively related to a variety of crime-related social control measures, such as rates of arrest (Mosher, 2001), resources and size of both law

enforcement (Chamlin, 1989), and rates of incarceration (Jacobs & Kleban, 2003). In addition to such formal social controls related to law enforcement, studies have shown that the perception of intergroup competition is related to negative attitudes toward people of color, lower levels of support for compensatory justice programs (i.e. affirmative action), and stereotyping (Riek, Mania & Gaertner, 2006; Zarate, Garcia, Garza, & Hitlan, 2004).

Negative attitudes toward people of color may also arise because of a perceived conflict of values or beliefs, even in the absence of the kind perceived realistic threats (economic, political, and criminal) hypothesized in early iterations of intergroup threat theory (Bobo, 1988; Kinder & Sears, 1981; McConahay, 1982). As Riek and colleagues (2006) note, "when the interests of a group as a whole are threatened, members perceive this as threatening even though self-interest is not directly impacted" (p. 337). Here, non-dominant racial/ethnic groups can be perceived to present a *symbolic threat* due to the belief that they violate values that are important to dominant groups (Sears, 1988; Stephen & Stephen, 2000). Symbolic threats are predicted to produce envy, anger and behaviors that diminish the accomplishments of people of color (Neuberg & Cottrell, 2002).

In the context of school discipline, both realistic and symbolic threats are relevant. For instance, symbolic threat may arise in predominantly-White school contexts if students of color are perceived to present a threat to the prevailing culture and norms within communities, schools, and classrooms (Rocque & Paternoster, 2011). While it is unlikely that students are perceived to present a traditional economic or political threat to school authorities, widely held stereotypes associating youth of color with criminality

(Blair, Judd, & Chapleau, 2004; Eberhardt, Goff, Purdie, & Davies, 2004) could also lead school authorities to perceive a form of realistic threat. In both cases, threat theories predict that officials may respond to their nascent feelings of fear toward people of color by using school discipline policies and practices to defend prevailing norms and maintain socio-cultural dominance (Payne & Welch, 2010; Welch & Payne, 2010).

Overview of Context and Methods

This study used a bivariate regression with controls to examine the relationship between the proportion of White students in a school and racial/ethnic disparities in suspension between White students and their counterparts from six Census-identified racial/ethnic groups (American Indian/Alaska Native, Asian, Black, Latinx, Native Hawaiian/Pacific Islander, Two or more Races) while controlling for several other factors related to school discipline outcomes. While the primary focus of this analysis was the main effect of the proportion of White students on the relative risk of suspension for student groups of color (as compared to White students). In order to isolate the effect of proportion White to the greatest extent possible, the analysis included control variables previously linked to differential outcomes in suspension rates. Likewise, the findings and interpretations also emphasized the main effect of the proportion of White students on the relative risk of suspension for student groups of color (as compared to White students). The presentatin of the effect of control variables within the model was limited to cases where findings diverge significantly from previous literature, or when effects related to the control variables support new or emerging findings within the school discipline literature.

This study used school level demographic and discipline data from almost 90,000 K-12 public schools in the U.S. Discipline data for this study was collected by the United States Department of Education as part of the Civil Rights Data Collection (CRDC) for the year 2013-2014. Data on student demographics, school size, student disability population(s), student poverty levels, and the characteristics of teachers and school professionals was obtained from the National Center for Education Statistics Common Core of Data.

The methodology was intended to provide greater context for understanding the relationship between the demographic composition of U.S. schools and the presence of racial/ethnic discipline disparities. Previous studies have documented how demographic composition affects the incidence and punitiveness of discipline for African American students (Skiba et al., 2014; Rocha & Hawes, 2009; Welch & Payne, 2010). Furthermore, research examining the magnitude of racial/ethnic discipline disparities has hinted that disparities between Black and White suspensions may be particularly salient in predominantly-White school contexts (Eitle & Eitle, 2004; Rausch & Skiba, 2006; Wallace et al., 2008). This study intends to weave the emphases in these lines of previous research together by using a) demographic composition (percentage of White students in a school) as the focal independent variable, and b) measurement of racial/ethnic disparities (relative risk ratios) as a dependent variable. By using relative risk ratios, the current analysis generated a metric describing the extent of disparity by comparing the rate at which discipline is applied to different racial/ethnic subgroups. The magnitude of racial/ethnic disparity is then correlated with the proportion of White students in schools, along with several other related variables. This kind of analysis can begin to provide

some level of prediction as to where students of color may have an elevated risk of being over-selected for suspension through racialized, overly-punitive discipline policies and practices.

Research Question

This study asks the following research question:

What effect does the proportion of White students in a school have on school-level risk of out of school suspensions for students of color as compared to White students?

It was hypothesized that the proportion of White students would have a significant positive relationship with the relative risk of suspension for students of color.

Several other variables that previous research has found to influence school discipline outcomes were included as control variables in the analysis, including (a) school gender balance, (b) the proportion of students with disabilities at school, (c) the proportion of English language learners, (d) the proportion of students eligible for free or reduced lunch, (e) the proportion of novice teachers, (g) the number of school counselors, (h) school configuration (elementary, middle, high), and (i) school locale (urban, suburban, rural, town). Broadly speaking, this analysis was intended to improve understanding of what kinds of schools exhibit large racial/ethnic gaps in suspension, which in turn has implications for interventions aimed at reducing racial/ethnic disparities in school discipline and the disproportionate harm of students of color that results from exclusionary discipline.

Significance Statement

Given the limited nature of previous research on how socio-demographic factors affect racial/ethnic discipline gaps (Eitle & Eitle, 2004; Freeman& Steidl, 2016; Thornton & Trent, 1988), this study provided much needed empirical evidence as to the kinds of school environments that put students of color at increased risk of exclusionary discipline (as compared to their White counterparts). Specifically, results showed that schools with higher proportions of White students, lower poverty levels, and elementary schools tend to exhibit larger racial/ethnic disparities in suspension. As the first known study to examine the discipline gap between White students and all U.S. Census designated student groups of color (American Indian/Alaska Native, Asian, Black, Latinx, Native Hawaiian/Pacific Islander, Two or more Races) this study also contributed to analysis of previously under-examined populations, including Native Hawaiian/Pacific Islanders and students of two or more races. This study drew upon large, national samples, which produced empirical evidence of previous undetected discipline disparities for Asian students while demonstrating substantial disparities across all student groups of color, suggesting that the problem of racial/ethnic gaps in suspension is likely worse than has previously appreciated. In addition, results provided insight into under-examined school contexts, such as predominantly-White schools and schools in locales historically associated with large proportions of White students, even as demographic trends indicate that students of color are rapidly dispersing into historically White school contexts (Colby & Ortman, 2014; U.S. Census Bureau, 2015). Further findings indicated that the variables that have best predicted the student-level incidence of discipline in past studies were not the same variables that best predicted school-level magnitude of racial/ethnic

disparities. Specifically, the analysis showed that attending predominantly-White schools may expose students of color to more racialized structures of school discipline. Therefore, results suggest that the kinds of interventions that may reduce the overall incidence or severity of discipline may be qualitatively different than those intended to reduce racial/ethnic gaps between Whites and students of color. The identification of what kinds of schools exhibit the largest discipline gaps can inform efforts intended to reduce the very real harm associated with the disproportional application of suspension to students of color.

Summary

This dissertation consists of five chapters. The current chapter provides a general introduction to the study. It presents initial evidence for the pervasive presence of racial/ethnic differences in school discipline and the negative consequences of disproportionately applied punitive discipline. Chapter Two presents an overview of the school discipline literature from the 1970s to the present. It includes research evidence related to the presence, consequences and predictors of disparities in school discipline. Chapter Two also presents CRT and ITT, the theoretical frameworks guiding this study. CRT and ITT provide a lens for interpreting pervasive and longstanding racial/ethnic discipline gaps in American schools, which function as site for the application of racialized policies, practices, beliefs at both the structural and individual level. Chapter Three outlines the study's methodology, which uses regression analysis to examine the relationship between the magnitude of racial/ethnic discipline disparities, the proportion of White students, and a number of control variables plausibly related to disproportional discipline outcomes. Chapter Four provides the results of the study used to answer the

research question. Chapter Five positions the results within the extant research literature through a presentation of key findings, their significance, and recommendations for policy, practice and future study.

Definition of Terms

The following key terms and concepts are defined in this study as follows:

School Discipline. School discipline typically refers to in-school suspension, out of school suspension, expulsion or referrals to law enforcement. This study examines out-of-school suspension only.

Racial/Ethnic Discipline Disparities. The proportion of discipline applied to a racial/ethnic subgroup within a school that is greater or less than the same racial/ethnic subgroup's proportion of the total school population.

Students of Color/People of Color: Those individuals not of White or European parentage who share the common experience of systemic racism (Alvarez, Liang & Neville, 2016). This study proceeds in full recognition of the fact that race is a social construct, but also that socially constructed notions of race have and continue to influence the beliefs, ideologies, and actions of individuals and social institutions.

School Authorities. School-based actors with the ability to define student behavior, initiate disciplinary action, or enforce punishment, including teachers, principals, administrators, staff, school resource officers, uniformed law enforcement officers or other actors with disciplinary authority.

Risk Index. A risk index (RI) (Hosp & Reschly, 2003; Reschly, 1997) is a method for indicating overrepresentation. It is the percentage of a group (i.e. American Indians) in a category (students suspended). In this study, RI is obtained by dividing the

number of times a particular discipline outcome (i.e. suspension) was applied to a given racial/ethnic group by the total number of students from that racial/ethnic group enrolled in a school. This yields an RI representing the number of instances of a particular punishment per 100 students in a given racial/ethnic group.

Relative Risk Ratio. The relative risk ratio (Coutinho & Oswald, 1998; Hosp & Reschly, 2003; Parrish, 2002) compares the rate at which discipline is applied to different racial/ethnic groups in order to generate a ratio describing the extent of discipline disparity. The RRR is generated by comparing the risk index for one group to the RI for another group.

Predominantly-White Schools. Predominantly white schools have a percentage of White students equal to or greater than 50 percent.

Racialized School Discipline. This term generally applies to discipline policies and practices that appear to be imbued with racial/ethnic stereotypes or negative perceptions toward students of color. It is also used in the context of empirical findings suggesting that disparate discipline outcomes could be reasonably attributed to racial/ethnic composition of place (schools, classrooms, communities).

Assumptions

This study makes the following assumptions:

1. The data reported by schools to the Office of Civil Rights is assumed to be a reasonably accurate portrayal of actual discipline outcomes in schools.
2. The demographic composition of schools (i.e. proportion of White students) is assumed to have some influence on the racial, ethnic and cultural dispositions of students, school authorities and other school-based actors.

3. School-level discipline outcomes are influenced by the beliefs, ideologies and dispositions of students, school authorities and other school-based actors.

CHAPTER 2: REVIEW OF LITERATURE

Overview

This chapter reviews the existing literature on school discipline disparities, the various factors that influence school discipline policy and practice, and the conceptual basis for interpreting the remarkable persistence of inequitable discipline outcomes. Current policies and practices are situated in the context of over four decades of a consistent and well-documented over-selection of students of color for exclusionary school discipline, even when variations in student behavior and socioeconomic status are accounted for. As the research suggests, it is school-level variables that are the most powerful predictors of discipline, particularly the racial/ethnic composition of schools, the attitudes, beliefs and dispositions of school authorities, and the presence of zero tolerance discipline policies. This chapter will also provide a synthesis of previous discipline research that indicates how the traits of schools with larger White populations (racial/ethnic balance, school personnel characteristics, and school wealth and resources) suggest that such schools constitute a potential site for particularly racialized forms of punishment. By documenting the factors that influence school discipline outcomes in concert with the traits of schools with larger White populations, this chapter seeks to construct a "profile" of these contexts as a site for the practice of school discipline within a social system characterized by the kinds of structural racism examined by critical race theory, as well as the individual dispositions emphasized by intergroup threat theories

Taken together, the extant literature provides ample ground upon which to base this study's use of racial/ethnic composition (proportion of White students) as an important predictor of the degree of discipline disproportionality in the context of schools with larger White populations.

Discipline Disparities

Who Experiences Disproportional Discipline?

As early as the 1970s, studies documented the overrepresentation of African American students in school suspension (Children's Defense Fund, 1975). Similar findings have emerged in relation to other forms of discipline, including office referrals (Skiba et al., 2002), expulsions (Kewel Ramani, Gilbertson, Fox & Provasnik, 2007), and referrals to law enforcement (Fabelo et al., 2011). In addition to African Americans, research also shows persistent discipline disparities for American Indians (Brown & Di Tillio, 2013; Krezmien et al., 2006; Rocque, 2010) and students with disabilities (Skiba et al., 2002; Wallace et al., 2008), and to a lesser degree, for Latinx populations (Peguero & Shekarkhar, 2011; Skiba et al., 2011), students with lower socioeconomic status (Christle, Nelson, & Jolivette, 2004; Raffaele Mendez et al., 2002; Skiba, Peterson & Williams, 1997), and LGBT youth (Himmelstien & Bruckner, 2011). It is worth reiterating that a quick glance through this list of student sub-groups reinforces the conclusion that disproportional discipline is applied most often and most consistently to marginalized social groups.

The Consequences of Disproportional Discipline

The overuse and inequitable application of exclusionary discipline has damaging effects on the academic outcomes, social/emotional well being and ultimately the life

chances of students that are most often the target of such punishments (Fabelo et al., 2011; Losen & Martinez, 2013; Wallace et. al., 2008).

Academic Achievement & Attainment. A growing body of literature has explored the ways discipline practices alternatively enable or constrain student academic outcomes (Annamma, Morrison & Jackson, 2014). Meta-analysis of 34 studies over a 26 year period reported a consistent inverse relationship between suspension and achievement, along with a significant positive relationship between suspension and dropout (Noltemeyer, Ward, & McLoughlin, 2015). Recent evidence shows that being suspended even once in ninth grade doubled the likelihood of dropping out, from 16% for those not suspended to 32% for those suspended just once (Balfanz, Byrnes & Fox, 2014). Lewis, Butler, Bonner III and Joubert (2010) studied a mid-western urban school district that enrolled 3,587 African American male students. They found that African American males were suspended for a total of 3,714 school days during one academic school year. This loss of classroom time correlated with lower proficiency on standardized test scores in reading, writing, science and math. The link between racial/ethnic differences in discipline and achievement has led some scholars to refer this relationship as "two sides of the same coin" (Gregory, Skiba & Noguera, 2010, p. 1).

Social & Emotional Factors. Exclusionary discipline and the attendant academic consequences also affect the social and emotional well-being of students. Suspension strongly correlates with future delinquency and substance abuse (Raffaele Mendez, 2003). Developmental research also indicates that frequent use of punitive discipline is at odds with the developmental challenges of adolescence and may inhibit social growth (American Psychiatric Association Zero Tolerance Task Force, 2008). Inappropriate use

of school discipline has also been linked to posttraumatic stress disorder, depression, anxiety, and aggressive behavior in and outside of school (Cameron & Sheppard, 2006). Yet, when they arise in relation to school disciplinary incidents, the signs of social and emotional damage are often "perceived as insubordinate or disruptive" (Edelman, 2007, p. 1), potentially evoking additional punishment and trauma while squandering the opportunity to re-examine punitive discipline in favor of more developmentally appropriate clinical and restorative interventions. Inequitable and overly punitive school discipline environments also undermine developmentally important teacher-student relationships (Hamre & Pianta, 2006; Losen & Martinez, 2013).

Juvenile Justice & the School-to-Prison Pipeline. By nature, suspension and expulsion remove students from class, and can also increase the amount of time that they spend unsupervised and with other out-of-school youth (Wallace et. al., 2008). The close relationship between exclusionary discipline and the criminal justice system (Nicholson-Crotty, Birchmeier & Valentine, 2009) has led scholars to note the increasing "criminalization of school discipline" (Hirschfield & Celinska, 2011, p. 1). Racial/ethnic and socioeconomic divisions amongst those most severely sanctioned in the juvenile justice system, the legal system and school discipline are remarkably similar (Advancement Project, 2005). For example, research showing that students of color are over three times more likely to be punished in school are directly mirrored in juvenile justice research, which has shown that in 1998 Latinx and African American youth were respectively three and six times more likely to be in jail as Whites for the same offenses (Poe-Yamagata & Jones, 2000). Furthermore, increased reliance on formal control measures (i.e. uniformed officers, surveillance systems, metal detectors, barred windows)

has created a prison-like atmosphere in many schools where students receive the message that they are apprentice criminals "preparing for prison" (Hirschfeld, 2008, p. 1; Watts & Erevelles, 2004). Many scholars have begun to refer to the glaring similarities between criminal justice policies and school discipline as the *school-to-prison pipeline* (Advancement Project, 2005; Wald & Losen, 2003).

Race, Ethnicity & School Discipline

While the process of school discipline is complex and determined by a number of interacting variables, race has been shown to be one of the best predictors of disciplinary involvement (Morrison & Skiba, 2001, Raffaele Mendez et al., 2002). Perhaps the most striking feature of the school discipline literature is the remarkable persistence and consistency of racial/ethnic discipline disproportionality. Studies have repeatedly found that African American students are three to six times more likely to be selected for exclusionary discipline (Children's Defense Fund, 1975; Eitle & Eitle, 2004; Rocque & Paternoster, 2011). The effect of students' race remains significant in large, multivariate studies that control for differences in (mis)behavior, socioeconomic status and a host of other relevant classroom and school variables (Rocque, 2010; Skiba et al., 2014; Welch & Payne, 2010). For example, Fabelo and colleagues (2011) conducted a longitudinal, multivariate study of *all* 7th grade students in Texas (with almost 5 million students) over the course of at least six years. The researchers controlled for 83 different variables (i.e. differences in student behavior, school discipline policies, socio-economic class, language proficiency, attendance, teacher experience/qualifications, and school resources/per-pupil expenditures) allowing them to isolate the effect of race alone on disciplinary actions. The study found that African American students were 31% more

likely to receive disciplinary action compared to otherwise identical White and Latinx students. Similarly, Skiba and colleagues (2011), studying 364 elementary and middle schools during the 2005-2006 school year, found that Black students were 2.19 (elementary) to 3.78 (middle) times more likely to be referred to the office as their White peers. While these results are concerning in their own right, the recognition that they closely echo the results from the four decades prior demonstrates that racial/ethnic disparities in school discipline have remained virtually unchanged.

Broadly speaking, the literature on racial/ethnic differences in school discipline has explored two categories of variables: student variables (i.e. race/ethnicity, misbehavior rates, SES) and school/classroom variables (i.e. discipline policies, principal disposition, teacher attitudes, school demographics). This division is indicative of the ongoing debate about the roots of racial/ethnic differences in school discipline. As Rocque (2010) explains, "(t)he key debate is whether disproportionate minority discipline is a function of differential behavior (Do students of color offend more frequently?) or a function of differential treatment (Are officials acting in a biased fashion when enforcing laws/rules?)" (p. 558). Thus, the distinction between student traits and non-student traits is important because it informs our understanding of where and with whom the predictors of racial/ethnic discipline disparities reside. In turn, this distinction has important implications for efforts to promote discipline equity.

Student Variables

In the face of pervasive and consistent discipline disparities, it may seem reasonable to hypothesize that more frequently disciplined students have identifiable, relevant traits that make them qualitatively different from their peers. This logic has

produced a large body of research examining the student traits that might contribute to racialized school discipline (Eitle & Eitle, 2004; Raffaele Mendez & Knoff, 2003; Reyes 2006; Skiba, 2000). The two most prominent hypotheses suggest that students of color a) differentially experience socioeconomic factors that collude with racial/ethnic differences (Skiba et al., 1997; Wu, Wu, Pink, Crain, & Moles, 1982), and/or b) exhibit differential behavior to produce a higher frequency of discipline (Hindelang, Hirschi & Weis, 1979; Murray & Herrnstein, 1994; Wilson & Herrnstein, 1985).

Socioeconomic Status. Student differences in socio economic status (SES) have been offered as an explanation for racial/ethnic discipline disparities. Here, it is hypothesized that African Americans may "disproportionately come from low-income households, may be overly represented among special education students, or may have missed more school than students of other races, which some researchers have correlated with misbehavior in school" (Fabelo et al., 2011, p. 46). Some scholars have also suggested the operation of an implicit "poverty hypothesis" for racial/ethnic disparities in school discipline, where low-income students of color are assumed to come from more difficult family and community contexts and thus more likely to violate school rules (Skiba, Eckes & Brown, 2010; Williams, 2015).

Numerous studies have found that low SES students are overrepresented in school discipline (Christle et al., 2004; Petras, Masyn, Buckley, Ialongo, & Kellam, 2011; Wu et al., 1982). Nichols (2004) found that students receiving free or reduced cost lunch (FRL) were suspended three times more often than students paying full price in a large Midwestern school district. Similarly, Hinojosa (2008) found that other factors associated

with SES, including father or mother absence and quality of home resources, were also predictors of the higher incidence of suspension.

While poverty levels clearly factor into disciplinary outcomes (see McCarthy & Hoge, 1987), the vast majority of studies have found that SES is limited in its ability to explain racial/ethnic differences in school discipline (Raffaele Mendez et al., 2002; Skiba et al., 1997; Welch & Payne, 2010). Using logistic regression to control for family structure, parental education, and socio-demographic factors related to students' region and urbanicity, Wallace and colleagues (2008) concluded that controlling for SES "reduces the magnitudes of the racial/ethnic differences only modestly, and all of the subgroups remain significantly different from their White counterparts" (p. 9). In short, student SES is a significant factor, but it is insufficient to explain racial/ethnic discipline disparities.

Behavioral Differences. Numerous school discipline studies have reported behavioral differences between student racial/ethnic subgroups (McCarthy & Hoge 1987; Skiba & Rausch 2006; Rocque, 2010; Wallace et al., 2008). In a study of a nationally representative sample of 8th, 10th and 12th graders, Wallace and colleagues (2008) found "relatively small" (p. 7) differences between White students and Latinxs (drug use, alcohol use, gun possession), Blacks (alcohol use, gun possession), and American Indians (gun possession). Roque's (2010) analysis of racial/ethnic differences in office referrals in 45 elementary school indicated that teachers reported more misbehavior from African American students on a student misbehavior scale. In contrast, McFadden, Marsh, Price & Hwang (1992) examined over 4000 discipline files and reported that while "generally speaking, there was little variability across the categories of rule violations…[W]hite

pupils had relatively high rates of truancy and safety violations and [B]lack pupils had lower rates of those same behaviors" (p. 143). Similarly, Shaw and Braden (1990) reported that White students were more often referred for severe rule violations than Black students. McCarthy and Hoge (1987) noted that the only two violations that showed significant black-white differences were skipping class and carving on desks, both of which were committed more often by White students.

Other studies have highlighted the presence (or absence) of racial/ethnic differences in student behavior depending on the relative a) subjectivity/objectivity or b) seriousness of the violation. Skiba and colleagues (2002) found no evidence that Black students received office referrals for either more types or more serious offenses. Rather, Black students were typically referred for subjectively defined offenses (i.e. disrespect, noisiness, threats, loitering) while White students were more often referred to the office for objective rule violations (i.e. smoking, vandalism, cutting class, obscene language). Likewise, Fabelo and colleagues (2011) conducted a longitudinal study of all students in Texas, and concluded that African American students were no more likely, and in many instances significantly less likely, than White and Latinx students to commit serious offenses that trigger mandatory disciplinary consequences. As the authors noted, "high rates of disciplinary involvement among African American students were driven chiefly by violations that are subject to the discretion of school employees" (Fabelo et al., 2011, p. 46).

Importantly, in the only known study to specifically examine predictors of student misbehavior (as opposed to the antecedents of disciplinary consequences), Finn, Fish and Scott (2008) found virtually no relationship between level of misbehavior and student

race or ethnicity at any degree of offense seriousness (low, intermediate, high). The only exception was the finding that Asian students committed significantly fewer intermediate violations.

As the literature above demonstrates, findings related to racial/ethnic differences in behavior are mixed and inconsistent. However, this body of literature (including studies that reported racial/ethnic differences) has been unanimous in the finding that differences in behavior cannot account for racial/ethnic disparities in school discipline (Wallace et al., 2008; Wu et al., 1982; Roque, 2010, Fabelo et al., 2011). As Kupchick & Ward (2011) have summarized, "there is a mountain of evidence" concluding that the relatively minor contribution of student behavior and class differences is dramatically eclipsed by the effect of race (Bradshaw, Mitchell, O'Brennan & Leaf, 2010; Horner, Fireman, & Wang, 2010; Kewel Ramani et al., 2007).

School and School Authorities Variables

Given that SES and behavioral variation have been unable to explain all (or even most) of the magnitude of racial/ethnic differences in school discipline, scholars have sought other explanations related to the traits of schools and school authorities. Wu and colleagues (1982) foreshadowed this shift in focus decades ago when they noted:

> To the extent that nonwhite minorities experience more suspensions than White students, after adjusting for their respective share of misbehavior, the additional suspension experienced by the nonwhite student is thus indicative of unequal treatment against them. The higher rate of suspension experienced by nonwhite minorities is therefore indicative of racial discrimination. (p. 40)

The clear implication is that differential *treatment* by schools and school authorities, not differences in student traits, is the primary driver of inequities in school discipline, particularly in the context of zero tolerance policies.

Zero tolerance policies. Since the 1990s, zero tolerance discipline policies have been studied extensively in the context of increasingly prevalent, punitive and racially disproportional U.S. schools. The term "zero tolerance" is not formally defined in the laws or regulations relevant to its establishment, however, the American Psychological Association Zero Tolerance Task Force (2006, p. 26) forwards the following definition:

Zero tolerance is a philosophy or policy that mandates the application of predetermined consequences, most often severe and punitive in nature, that are intended to be applied regardless of the apparent severity of the behavior, mitigating circumstances, or situational context.

Zero tolerance in education emerged alongside "get-tough" national criminal policies during the 1990s, such as habitual offender statutes, minimum sentencing guidelines and three strikes laws of the time that were intended to prevent crime and violence through increasing punitive consequences (Casella, 2003; Hirschfeld & Celinska, 2011). Over time, the mandate of zero tolerance was expanded, first to weapons, and criminal activity related to drugs, alcohol, tobacco, threats and swearing (Skiba, 2000), but subsequently to a broad array of relatively ordinary student behaviors not traditionally understood as criminal, such are insubordination, disrespect, or disruption that are generally unrelated to school safety (Advancement Project, 2000; Gregory et al., 2010). The expansion of zero tolerance policies was mirrored in other forms of social and behavioral control in schools, including dress codes, student ID

badges, surveillance systems, metal detectors, uniformed resource officers, locker searches and gun-sniffing dogs (Welch & Payne, 2010).

Under zero tolerance, harsh discipline and office referrals dramatically increased as suspension became the punishment of choice under zero tolerance policies (Kewel Ramani et al., 2007; Losen & Martinez, 2013; Wallace et al., 2008). Zero tolerance also contributed to larger racial/ethnic discipline gaps as school authorities applied the no-nonsense ethos to a growing array of student behaviors. Research has found that students of color are often sent to the office based on subjective offenses (Gregory et. al., 2010), and that racial/ethnic discipline disparities are more likely to be found in relatively minor, subjective offense categories (Losen & Martinez, 2013). Skiba and colleagues (2002) conducted a study of middle-school students in a large, urban mid-western public school district located in one of the 15 largest cities in the United States. Drawing on the discipline records of 11,001 students in 19 middle schools, they concluded that students of color were disproportionately referred for subjective offenses (e.g. loitering, disrespect, excessive noise), while White students tended to be referred for objective violations (e.g., smoking, vandalism, obscene language).

Unfortunately, suspension does not appear to deter future misbehavior (American Psychiatric Association Zero Tolerance Task Force, 2008). Raffaele Mendez (2003) conducted a longitudinal study of suspension in a Florida school system in which suspension rates were collected for student cohorts in grades 4 through 12. Despite a rapid decline in enrollment in grades 9-12, suspension rates remained between 18-20%, suggesting that out-of-school suspension does not work as a deterrent of misbehavior for suspended students or their peers. McFadden and colleagues (1992) conducted a one-year

investigation in Florida, which found that 25% of the students receiving suspension or corporal punishment committed more than five disciplinary offenses, 75% of the students committed between one and five offenses, and less than 1% of the students committed only one offense. Zero tolerance also fails to reduce violence or increase safety in schools (Advancement Project, 2000). Skiba and Peterson (1999) used National Center for Education Statistics data on school violence to show that, after four years of implementation, schools that use zero tolerance policies are still less safe than those without such policies.

Through their reliance on suspension and expulsion, zero tolerance policies have had particularly damaging effects on the academic attainment and the social/emotional well-being of the urban students of color that are most often the target of such punishments (Losen & Martinez, 2013; Wallace et. al., 2008; Welch & Payne, 2010). Furthermore, the close relationship between zero tolerance and the criminal justice system, often referred to as the school-to-prison pipeline (Advancement Project, 2005), has led scholars to examine the link between the "criminalization of school discipline" (Hirschfield & Celinska, 2011, p. 1) and patterns of overrepresentation of youth of color in juvenile and adult prisons (Annamma, Morrison & Jackson, 2014).

There are some indications that zero tolerance is falling out of favor amongst educational policy makers. In 2014, the US Secretary of Education Arne Duncan and U.S. Attorney General Eric Holder released guidelines urging schools to abandon zero tolerance discipline policies (U.S. Department of Education, 2014). Officials point to evidence that students of color tend to be disciplined more harshly and more frequently than their White counterparts under zero tolerance. Legal and educational challenges to

zero tolerance have been mounting for the past decade (St. George, 2011), alongside civil rights investigations related to racial/ethnic disparities in discipline conducted by the U.S. Departments of Education and Justice (Department of Education, 2014; U.S. Department of Education, 2016a). However, it appears doubtful that such oversight will continue in the administration of Donald Trump (Green, 2017).

The attitudes & dispositions of school authorities. Discipline studies have for decades invoked the presence of racial/ethnic biases (implicit and explicit), negative stereotypes and discriminatory attitudes to help explain racial/ethnic differences in discipline outcomes (McFadden et al., 1992; Rocque & Paternoster, 2011; Wallace et al., 2008). Scholars from a variety of disciplines and other areas within education have suggested that pervasive cultural attitudes and social arrangements are related to differential outcomes in education, political economy, legal systems and society at large (Apple, 1990; Delpit, 2006; Steele, 1997).

Given that most disciplinary action begins in the classroom (Skiba et al., 2002), the characteristics of teachers and their dispositions toward students and classroom management are of central importance in examining racial/ethnic discipline disparities (Ferguson, 2001; Payne & Welch, 2010; Townsend, 2000). Over the past 15 years many studies have sought to examine the "close relationship between teachers, the racial/ethnic beliefs and values teachers bring into the classroom, their pedagogy, and the outcomes of their students" (Williams, 2015, p. 13).

Hinojosa & Moras (2009) found that teachers were significantly more likely to exhibit racial/ethnic attitudes that are less tolerant than those of similar education levels in other fields. Furthermore, research indicates that discipline outcomes vary significantly

across and between teachers and classrooms (Gregory & Weinstein, 2008; Gregory & Thompson, 2010). There is mounting empirical evidence that questionable attitudes and dispositions are a factor into the production of racial/ethnic variation in discipline outcomes. Research has suggested that Black students are frequently stereotyped as "troublemakers" (Fenning & Rose, 2007; Okonofua & Eberhardt, 2015). Furthermore, research has suggested that the cultural characteristics (e.g., dress, speech patterns, patterns of social interaction) of students of color can lead school authorities to perceive them as having more negative demeanors, a longer history of misbehavior, and lower grades than White students (Neal et al., 2003).

Studies have documented how attitudes and dispositions are played out in school discipline practices. Bradshaw and colleagues (2010) reported that teachers gave Black students more disciplinary referrals even after controlling for the same teachers' ratings of classroom behavior. Similarly, a 2010 study found that among students who were classified by teachers as overtly aggressive, African Americans were more likely to be disciplined than any other group (Horner et al., 2010). In other words, the frequency of teacher's disciplinary actions against students of color went well beyond even what their own independent, explicit reports of student behavior would predict.

Teachers' attitudes and dispositions also appear to influence the escalation and severity of discipline. In a longitudinal qualitative analysis of classroom observations, videotaped lessons, and interviews, Vavrus and Cole (2002) found that Black students were most often suspended when school authorities "singled out" a particular "disruptive act among many" (p. 1). As a result, nonviolent events were escalated to suspension in the absence of any clear rule violation. This led the researchers to conclude "that

removing a student from class is a highly contextualized decision based on subtle race…relations that cannot be adequately addressed in school discipline policies" (Vavrus & Cole, 2002, p. 1). In a study of teachers' implicit biases (those that occur outside of conscious awareness), Okonofua & Eberhardt (2015) found that teachers felt Black students' misbehavior should be met with more severe discipline even when Black and White students behave in the same manner. As the authors summarized, "We have shown experimentally, for the first time, that teacher responses can contribute to racial/ethnic disparities in discipline…[and] may even help to drive racial/ethnic differences in student behavior" (Okonofua & Eberhardt, 2015, p. 622).

Further evidence of the effects of teachers' attitudes and dispositions can be found in studies that have explored potential ways to attenuate racial/ethnic disparities in school discipline. Researchers have found that when otherwise similar Black and White students are placed with same-race teachers, the classroom behavior of Black students was rated more favorably than was the behavior of White students (Downey & Pribesh, 2004). In recognition of the fact that discipline is often bound up in reciprocal relationships between students and teachers, Gregory and Weinstein (2008) explored the intersection of teachers' attitudes, students' attitudes and office referrals in a two-part, mixed-methods study of a large urban school district. The researchers found that teachers who referred a student rated the same student lower than teachers who students considered trustworthy and caring. Referring teachers were perceived by their students as less engaged than caring teachers, and self-reports indicated that students were less resistant with caring teachers. In similar research, Gregory and Thompson (2008) documented more variance *within* the disciplinary experience of individual student experience than

between students. In other words, the same students tended to have positive and supportive relationships with some teachers, and contentious relationships with others depending on the attitudes and perceptions of students and teachers. While they have been the subject of much less empirical research, the attitudes and dispositions of preservice teachers also appear to exhibit beliefs and dispositions that may contribute to racial/ethnic discipline disparities. Studied have found that the implicit attitudes (Glock, 2016) and disciplinary intervention strategies (Glock & Karbach, 2015) chosen by preservice teachers exhibited pro-White bias and racial/ethnic stereotypes, which resulted in harsher discipline for students of color. Glock and Karbach (2015) noted that school authorities who hold favorable attitudes toward racial majority students and negative attitudes toward students of color "might contribute to disadvantages experienced by racial minority students" (p. 59).

As discipline escalates out of the classroom to the front office, the beliefs and attitudes of administrators also become relevant. In a study of middle schools, Skiba and colleagues (1997) found significant variation between the seriousness of infractions and the severity of punishment given by administrators. As was the case with teachers, differing administrator dispositions appear to produce considerable inconsistency in discipline outcomes (Advancement Project, 2000; Fenning & Rose, 2007). Mukuria (2002) found that principal attitudes varied widely between schools with higher numbers of suspensions for African American students then those with lower numbers. Principals who were willing to consider student contextual factors and exhibited a clear philosophy for school discipline were less likely to resort to suspension than those principals that tended to strictly follow disciplinary policy without attending to context. Studies also

indicate that administrators are more likely to suspend or expel Black students when they commit the same kinds of infractions as White students (Skiba et al., 2011; Nicholson-Crotty et al., 2009). In addition, Skiba and colleagues (2014) found that principal perspective on discipline was predictive of disciplinary practices concerning the use of out-of-school suspension and expulsion. Specifically, "schools in which principals expressed attitudes more favorable toward school exclusion, students were significantly more likely to receive out-of-school suspension and expulsion relative to in-school suspension" (Skiba, et al., 2014, p. 657).

Demographic composition. One of the most striking findings of recent discipline research is that that the mere presence of large numbers of Black students contributes to a higher incidence of suspension and harsher discipline practices. Welch and Payne (2010) used a sample of over 800 schools to test of the effects of racial/ethnic composition on discipline practice. They found that schools are more punitive and less restorative when there are more black students enrolled in them regardless of the amount of school misbehavior, student economic disadvantage, school urbanicity or training of faculty and administration (Welch & Payne, 2010). Rocque and Paternoster (2011) showed that elementary schools with a higher proportion of Black students have higher mean levels of office referrals even when controlling for a variety of school, teacher and student characteristics. In one of the only studies to simultaneously examine student level variables (i.e. race, misbehavior rates, class) and the school level variables (i.e. principal disposition, % Black enrollment) Skiba and colleagues (2014) identified the relative contribution of variables on multiple levels. Like many previous studies, the researchers found that Black students were more likely to be suspended after controlling for the

effects of infraction type and class variables. However, when school level variables including the percentage of Black students and principal attitudes were introduced into the analysis, the effect of individual student race on the likelihood of suspension was reduced to non-significance (Skiba et al., 2014). This is the only known study to fully account for racial/ethnic disparities in school discipline.

The clear implication is that, while individual-level racial/ethnic bias most certainly affects discipline outcomes, the characteristics and dispositions of schools and school authorities constitute a *structural system* that functions to disproportionately select marginalized social groups for harsh discipline regardless of variation amongst students.

Discipline in Predominantly-White Schools

While scholars have noted that schools share many of the same challenges across locales (Milner, 2012), clear divisions exist between schools' geographic location, demographic make-up, and socioeconomic characteristics (U.S. Department of Education, 2010a). As previous sections have suggested, these divisions also represent significant factors in school discipline research (Finn et al., 2008; Skiba et al., 2002; Wallace et al., 2008; Wu et al., 1982). This study recognizes that these relationships are fluid and may not hold true indefinitely. Demographic trends project growing numbers of students of color in U.S. schools in the next several decades (Colby & Ortman, 2014; U.S. Census Bureau, 2015), and the proportion of students of color is also rising in historically White communities with large numbers of predominantly-White schools (Orfield & Frankenberg, 2014).

The research literature contains no known studies explicitly addressing racial/ethnic discipline disparities in predominantly-White school communities.

However, close reading of past studies examining race/ethnicity and school discipline foregrounds how schools with large proportions of White populations and those without are constituted in fundamentally different ways that may influence how discipline is applied to students of color based on differing school demographic composition.

Meanwhile, a tremendous amount of research has explored how both student-level and school-level factors contribute to disproportional discipline of students of color more generally. As previously noted, four decades of increasingly sophisticated research has reached the unanimous conclusion that racial/ethnic disproportionality remains in school discipline even after controlling for a host of other variables. Furthermore, the most recent multivariate studies strongly suggest that school-level characteristics account for a much greater proportion of racial/ethnic discipline disparities than do student-level characteristics (Fabelo et al., 2011; Rocque & Paternoster, 2011; Skiba et al., 2014, Welch & Payne, 2010). This section highlights three categories of school-level factors that may influence the differential disciplinary experiences of students of color attending schools with larger White populations. When analyzed in concert, these three categories suggest that students of color attending schools with larger White populations must navigate a unique and perhaps more perilous brand of discipline inequality than their urban counterparts.

Racial Balance

Research has shown that a higher percentage of students of color (especially Black students) is one of the strongest predictors of racial/ethnic disproportionality in exclusionary discipline and the incidence of office referrals net of other school and individual level factors (Rocque & Paternoster, 2011; Skiba et al., 2013, Welch & Payne,

2010). Specifically, research focused on the percentage of Black students has suggested that growing numbers of students of color will result in more punitive, racially-motivated discipline based on perceptions of racial threat (Welch & Payne, 2010). Rocque and Paternoster (2011), in a study that found evidence of racial threat on the elementary school level, suggested that "as the black student population increases, teachers may perceive black student misconduct differently, as perhaps more menacing or more of a threat to their control, and respond to such conduct by African Americans more punitively" (p. 655). Studies have also shown that more racially balanced schools have lower levels of racial/ethnic discipline disproportionality (Eitle & Eitle, 2004). Put differently, students of color attending highly segregated schools with larger White populations may be at higher risk for discipline when compared to White students in the same schools. Eitle & Eitle (2004) also call attention to studies from the era of desegregation (1970s and 1980s), which found higher incidence and greater disproportionality in Black suspension rates in formerly White schools.

Racial/ethnic balance interacts with levels of student delinquency in ways that raise additional concerns for students of color in schools with larger White populations. Welch & Payne (2010) found that the percentage of Black students had a stronger effect on harsh discipline practices in schools with less delinquency and drug use. NCES data shows that schools associated with higher proportions of White students experience lower incidence of crime than more diverse schools (U.S. Department of Education, 2006). Given that Whites make up large majorities of public school student populations in suburban, rural and town locales (Snyder, de Brey, & Dillow, 2016), this research suggests that schools in less diverse locales may be more likely to punish perceived

student misbehavior based on the racial/ethnic composition of the student body, and that disciplinary actions against students of color in schools with larger White populations may be more punitive in nature (Welch & Payne, 2010).

While previous scholarship makes it clear that demographic composition affects discipline in schools with large and growing populations of students of color, these populations are growing in historically White schools with higher proportions of White students. Generally, White students have decreased (-8%) in U.S. public school enrollment whereas students of color increased (+10%) (U.S. Department of Education, 2007; U.S. Department of Education, 2010). Furthermore, students of color are no longer concentrated only in urban centers and the South. Students of color are becoming more dispersed, with increasing concentrations in many historically White communities across the nation (Orfield & Frankenberg, 2014). Thus, schools with larger White populations must navigate the demographic reality of rising populations of color with the associated challenges of implementing equitable disciplinary strategies that effectively deal with students from a variety of cultural backgrounds. It is plausible that students of color in schools with larger White populations might face the same kind of more punitive, racially-motivated discipline documented in more diverse school contexts. However, given the lack of research on discipline in schools with larger White populations, this question remains unclear.

Teacher Characteristics

Teachers play a critical role in school discipline. Because most discipline referrals originate in the classroom (Skiba et al., 2002), the characteristics of teachers and their dispositions toward students are of central importance (Eitle & Eitle, 2004; Gregory &

Thompson, 2010; Payne & Welch, 2010; Raffaele Mendez et al., 2002; Rocque & Paternoster, 2011). Research findings indicate that teacher race/ethnicity and student-teacher race/ethnicity matching influence the harshness and incidence of school discipline (Horner et al., 2010). According to the NCES, 84.6 % of teachers in historically White schools are White, 6.3% are Black, 6.2% are Latinx, 1.4% are Asian and 0.4 % are American Indian, Alaska natives, Hawaiian natives or Pacific Islanders. This compares with 71% White, 12% Black, 13.1% Latinx, 2.2% Asian and 0.7% American Indian, Alaska natives, Hawaiian natives or Pacific Islanders in more diverse, urban school contexts (U.S. Department of Education, 2009a). These data suggest that less diverse teaching corps may disproportionately contribute to racial/ethnic discipline disparities in schools with larger proportions of White students (Eitle & Eitle, 2004).

Teacher experience and qualifications also influence discipline. Several studies have reported that more experienced teachers produce higher suspension levels for students of color after statistically controlling for various school-level variables such as racial/ethnic balance, poverty measures, student homogeneity (Eitle & Eitle, 2004; Raffaele Mendez et al., 2002). Government statistics demonstrate historically White schools tend to have more experienced teachers (defined as over four years) (U.S. Department of Education, 2009b). More diverse schools staff the highest percentage of novice (< 4 years) teachers (U.S. Department of Education, 2009b). Mirroring findings on teacher education levels, research suggests that schools with more White students have more highly qualified teachers as defined by those with more experience, educational attainment, and teaching certifications (Jacob, 2007; Lankford, Loeb & Wyckoff, 2002). Several studies have found an association between more highly

qualified teachers and the over-selection of students of color in both suspension and office referrals (Eitle & Eitle, 2004; Rocque & Paternoster, 2011).

Research indicating that more experienced, qualified teachers disproportionately contribute to racial/ethnic discipline disparities may suggest that these teachers have had more time in the classroom to develop and/or reinforce deficit theories and negative stereotypes of students of color, which are expressed through their discipline practices. Also, because a higher proportion of more experienced, qualified teachers practice in historically White schools with greater racial/ethnic imbalance, they may have increased perceptions of criminality and intergroup threat toward students of color. Indeed, research suggests that racial/ethnic (im)balance affects teacher perceptions. Payne & Welch (2010) found that increasing percentage of Black students was associated with higher teacher perceived lack of safety. The researchers have also noted that schools with higher proportions of White students appear to be less likely to have discipline training for teachers (Payne & Welch, 2010). Thus, while the literature concludes that teacher characteristics in schools with larger White populations (with more experienced, qualified teachers) may be more likely to produce racial/ethnic discipline disparities, those same teachers may be less likely to have access to the types of discipline training that have the potential to mitigate the effects of cultural differences and implicit bias on school discipline.

School Wealth & Resources

While there is a lack of available data on differences in school wealth and resources specifically between schools with larger White populations and those without, metrics do exist based on school locale (urban, suburban, rural, town). As stated

previously, schools with larger White populations are positioned predominantly (though not exclusively) in suburban, rural and town locales. These locales have per pupil expenditures exceeding those in more diverse schools (Betts et al., 2000; U.S. Department of Education, 2015). NCES reports also indicate that teachers in urban districts have fewer resources at their disposal (U.S. Department of Education, 1996). Resources, along with previously mentioned differences in racial/ethnic balance and teacher characteristics, also influence school disciplinary disposition and behavior control strategies (Fenning & Rose, 2007; Hirschfield, 2008). Indeed, metal detectors, gates and barricades are more common in racially/ethnically diverse, urban, schools, while drug sniffing dogs are more common in predominantly-White schools (DeVoe, Peter, Noonan, Snyder & Baum, 2005; Gottfredson et al., 2000). Such findings led Hirschfield (2008) to note that "the gated community may be a more apt metaphor to describe the security transformation of affluent schools, while the prison metaphor better suits that of inner-city schools."

Recent research sheds light on the intersection of school resources and disciplinary action. Skiba et al. (2014) found that schools with lower levels of student poverty (as proxied by free and reduced lunch eligibility) have higher rates of expulsion, perhaps the most punitive of disciplinary measures. Since these schools tend to have fewer students of color and less student poverty, this finding is consistent with previously mentioned research demonstrating that disciplinary actions against students of color in schools with larger White populations may be overly punitive in nature based on the operation of racial threat and pervasive negative stereotypes (Welch & Payne, 2010). However, previous research came to a seemingly contradictory conclusion; schools with

higher proportions of White students are more likely to offer less punitive and more restorative disciplinary actions such as referrals to counselors and treatment centers (Fenning, Wilczynski & Parraga, 2000).

To elucidate this seeming contradiction, is important to note that such restorative practices are resources intensive and often require a team of school-based professionals with expert behavioral consultants (Fenning & Rose, 2007). One way that levels of school resources and wealth are expressed is in differential staffing of school professionals (i.e. counselors, psychologists, social workers). NCES data indicates that schools with larger White populations generally staff more school professionals, especially counselors and psychologists (U.S. Department of Education, 2009c). It is reasonable to suppose that greater resources and increased staffing of school professionals enables Whiter schools to offer clinical services that are not available in less affluent, more racially/ethnically diverse contexts. Considering that schools with larger White populations can simultaneously display higher levels of extremely punitive and more restorative discipline practices (i.e. referrals to medical professionals), the question then arises: Which types of students within schools with larger White populations are being provided access to clinical interventions that tend to be less punitive and more restorative? Previous research suggests that schools with larger White populations may tend to prescribe harsh, punitive punishments to students of color while reserving more clinical interventions for White students (Payne & Welch, 2010; Welch & Payne, 2010).

Reform and Intervention

A growing body of literature on alternative approaches has outlined several interventions aimed at reducing the racial/ethnic discipline gap. Research has identified several empirically tested alternatives that can effectively address student misbehavior and simultaneously have positive effects on educational outcomes and student growth (Gregory et. al., 2010; Losen & Martinez, 2013, Richart, 2004). Schoolwide Positive Behavior Intervention Support (SWPBIS) is the most widely used and well-researched alternative to punitive discipline policies that tend to over-select students of color for punishment. SWPBIS espouses the proactive teaching of expected behaviors and development of positive teacher–student relationships while promoting a better school climate for all students (Fenning & Rose, 2007). Numerous studies have shown that SWPBIS has the potential to limit disciplinary incidents and improve student outcomes within urban school contexts (Bradshaw, Mitchell & Leaf, 2010; Horner et al., 2009; Warren et al., 2003). Research also suggests that schools and districts can reduce both suspensions and racial/ethnic disparities more effectively if they revise their school codes to align with the positive and constructive framework of SWPBIS (Fenning et al., 2013). Lassen, Steele, and Sailor (2006) studied a school-wide SWPBIS implementation in an urban, inner-city middle school in the Midwest over a three-year period. Data on referrals, suspensions, standardized test scores, and treatment fidelity were gathered and analyzed. Researchers found significant reductions in referrals and suspensions and increases in standardized math and reading scores (Lassen et al., 2006).

Restorative justice models and policies (Costello, Wachtel, & Wachtel, 2010; Gregory, Clawson, Davis, & Gerewitz, 2016; Schiff, 2013) may also help address the

racial/ethnic discipline gap. Restorative models are typically schoolwide programs that aim to transform students' interactions with peers and school authorities using community-building activities in classrooms (circles) and a relationship-based process to resolve disputes (conferences) (Gregory, Bell, & Pollock, 2014). Restorative circles and conferences include teachers, students and a third-party facilitator, and allow for responsibility for disciplinary incidents to be shared between students and school authorities. They can also help highlight how disciplinary incidents are escalated by external factors, like the stress of teaching large classes or the application of bias and stereotypes in interpreting student behavior. Research has found that restorative practices can decrease the use of punitive school discipline (Lewis, 2009), more positive relationships between teachers and their diverse students, and fewer discipline referrals for Latinx and African American (Gregory et al., 2016).

Recent scholarship from leading discipline scholars has leveraged aspects of SWPBIS and restorative practices to produce an intervention framework intended to eliminate disparities in school discipline (Gregory, Skiba, & Mediratta, 2017, p. 268). The framework includes ten principles, including supportive relationships, opportunities for social learning, a problem-solving approach, inclusion of student and parent voice, and provisions for the reintegration of students after discipline. Along with these aspects drawn from SWPBIS and restorative practices, this work calls upon school authorities practice culturally relevant practices and to "explicitly consider" issues of culture, race, power, and privilege in addressing inequitable discipline outcomes (Gregory et al., 2017, p. 268). It also advocates for data-based inquiry into policies, practices and school contexts that appear to produce inequitable discipline outcomes. However, as Gregory

and colleagues (2017) note, "there is insufficient empirical evidence to indicate which combination of the 10 principles from the Framework should be implemented together, or which principles might be prioritized over others to reduce gender and race disparities in school discipline" (p. 271).

Theoretical Framework

Much of the empirical literature on racial/ethnic disparities in school discipline appeals to (but does not directly study) notions of cultural differences, negative attitudes and deficit ideologies as a contributor to differential outcomes. While such empirical works rarely moves far beyond demonstrating the influence of different variables, a smaller body of theoretical scholarship has attempted to explain *how* and *why* some students are consistently over-selected for school discipline. Given its prominence as a predictor of disparities, it is not surprising that theoretical works have focused primarily on race to frame their explanations. This section provides an overview of the ways that Critical Race Theory and Intergroup Threat Theory can inform understanding of racial/ethnic discipline disparities in educational institutions. Also, the inclusion of both theoretical traditions allows for a more nuanced treatment of the ways that race/ethnicity may influence discipline policy and practice. Specifically, it allows for the consideration of both macro-level, structural racism using CRT, as well as the micro-level, individual dispositions and activities of school personnel that have the potential to contribute to differential discipline outcomes between racial/ethnic student groups.

Critical Race Theory

Critical Race Theory (CRT) emerged from legal scholarship in the 1970s as a critique of the ways that law and the legal system contribute to the oppression of students

of color (Bell, 1973). It has since been applied to challenge the legitimacy of the social institutions (property rights, the construct of race, public education) as well as a host of venerable social ideals, such as the "rule of law," "equal opportunity," and "equal protection" (Ladson-Billings & Tate, 1995; Watts & Erevelles, p. 275). While purposive individual acts of discrimination are an important component of racism, CRT focuses on how fundamental *structures,* such as legal and property rights, function to preserve White control over power and material resources in society (Parker & Lynn, 2002; Sleeter, 2012).

In education, CRT positions racism as the central construct in understanding inequitable school outcomes (Ladson-Billings & Tate, 1995). Critical race scholars argue that historical realities and current material structures function to deny students of color the full "use and enjoyment" of education and other social structures (Ladson-Billings & Tate, 1995, p. 59; Sleeter, 2012). Based on the understanding that "inequalities are a logical and predictable result of a society characterized by institutional racism" (Ladson-Billings and Tate, 1995, p. 47), CRT rejects the purported race neutrality of school discipline policies and problematizes the notion of "fairness" in schools (Watts & Erevelles, 2004). The "ordinariness" of structural racism (Ford & Airhihenbuwa, 2010, p. 33) produces a jarring dissonance between the egalitarian promise of education as the "great equalizer" (Mann, 1849) and the reality of persistent racial/ethnic disparities in discipline, achievement and school funding/resources. Thus, for CRT scholars, schools represent an important structural context where students are rewarded for displaying White cultural norms and sanctioned for cultural practices (e.g., dress, speech patterns,

patterns of social interaction) associated with other cultural groups. (Gay, 2006; Neal et al., 2003; Townsend, 2000).

Intergroup Threat Theories

Intergroup threat theories (Blalock, 1967; Blumer, 1958; Stephan & Stephan, 1996) suggest that increasing numbers of individuals of color within a bounded system, such as a school, can be perceived as a criminal, social or cultural threat by dominant groups, resulting in prejudice, negative stereotypes and increased social control measures. Early iterations of group threat suggested that increasing numbers of people of color are perceived to present an economic, political or criminal threat to dominant social groups (Blalock, 1967; Sherif & Sherif, 1969). The perception of such *realistic* threats presumably produced negative attitudes toward people of color and led to the imposition of punitive social controls to maintain social dominance. Research in legal studies and criminal justice has illustrated the operation of intergroup threat by demonstrating that increasing concentrations of people of color within a bounded system (i.e., a neighborhood or school) are consistently and positively related to a variety of crime-related social control measures, such as rates of arrest (Mosher, 2001), resources and size of law enforcement (Chamlin, 1989), and rates of incarceration (Jacobs & Kleban, 2003). Studies have also shown that the perception of intergroup competition is related to negative attitudes toward people of color, lower levels of support for compensatory justice programs (i.e. affirmative action), and stereotyping (Riek et al., 2006; Zarate et al., 2004).

Subsequent iterations of intergroup threat emphasize how groups may present *symbolic* threats to dominant social groups (Stephan & Stephan, 1996). Symbolic threats

are perceived as a threat to cultural values, moral, beliefs or worldview (Riek et al., 2006), and produce negative dispositions and behaviors that diminish the accomplishments of marginalized social groups (Neuberg & Cottrell, 2002). Empirical tests of symbolic threat have found that while Blacks were more likely to elicit emotions of fear and anger, groups such as Native Americans produced different reactions of pity and guilt (Neuberg & Cottrell, 2002). Such findings speak to the complexity and nuanced nature of intergroup threat, where perceptions of threat may produce qualitatively different emotional and behavioral responses toward different groups. For example, social control measures directed toward Latinxs based on perceived threats related to nationality and immigration status may differ from measures directed at Asians based on perceptions of intellectual superiority, which in turn may differ still from reactions to Black youth based on perceived threat of criminality.

The kinds of reactions observed and predicted in the intergroup threat models discussed above have also been found in education. Welch and Payne (2010) found that high concentrations of Black students predicted the 1) presence of punitive discipline policies, 2) higher levels of harsh discipline and 3) fewer restorative discipline measures (i.e. referrals to counseling or treatment centers). Similarly, Rocque and Paternoster (2011) found that elementary schools with higher proportions of African American students have significantly higher levels of disciplinary referrals regardless of both classroom/teacher and individual student characteristics. Thus, when applied to racial/ethnic differences in school discipline, intergroup threat illuminates how White cultural norms in the classroom, combined with negative attitudes and dispositions may lead school authorities to interpret the conduct of students of color as a symbolic to their

individual or school values, eliciting a more punitive response to the conduct of Black students than might be applied to identical behavior by White students (Rocque & Paternoster, 2011).

Summary & Conclusion

Despite the well-documented harm done by the disproportionate application of punitive discipline, students of color are consistently disciplined more often and more harshly than their White counterparts. While students' behavioral and socioeconomic traits are (at times) predictors of exclusionary discipline, the traits of students have proven insufficient to explain racial/ethnic differences in school discipline. Rather, the characteristics of schools and school authorities, particularly racial/ethnic dispositions and racial/ethnic composition, appear to be the most powerful predictors of inequitable discipline.

Schools' geographic location, demographic make-up and socioeconomic characteristics also are related to discipline outcomes. This review has highlighted several characteristics of schools with larger White populations that raise the potential for higher levels of racialized discipline and higher magnitudes of racial/ethnic disproportionality. These characteristics include racial/ethnic composition, teacher characteristics and metrics of school wealth and resources. Meanwhile, demographic trends indicate that there will be increasing school-based interactions between predominantly-White school personnel and increasingly diverse student populations. Moreover, increasing numbers of students of color are attending schools with larger White populations. Discipline outcomes in schools with larger White populations, as conditioned by prevailing policies and practices, provide a conceptual link to the culture of discipline operating in

classrooms and at schools. The demonstrated importance of school's racial/ethnic student composition provides grounds for this study's focus on the percentage of White students as a plausible predictor of discipline disparities in schools with larger White populations. Yet, little empirical research has been conducted on the discipline outcomes of students of color in these environments.

To the degree that schools with larger White populations exhibit a culture based on the values, beliefs and norms historically associated with White culture, threat theories provide insight regarding the motivation behind and operation of discriminatory discipline practices. Further, it provides a theoretical basis for the assumption that students of color may face increased risk of racially motivated disciplinary policies and practices in schools with larger White populations. In ITT discipline becomes a means to diminish the position of students of color within prevailing social system, to deny students of color the full benefit of education, thus allowing dominant cultural groups to maintain control of classrooms, schools, the political economy and larger society. To the extent that racial/ethnic discipline disparities are a consequence of systemic racism and the attendant structural conditions that deny students of color the full benefit of education and other social institutions, CRT challenges the legitimacy of school discipline as it is currently practiced, and rejects the prevailing discourse that positions the practice of education as race-neutral, egalitarian and meritocratic.

CHAPTER 3: METHODOLOGY

The purpose of this study was to examine the relationship between racial/ethnic discipline gap (between Whites and students of color) in suspension and the proportion of White students in U.S. public schools. The analysis controlled for the influence of several other school-level characteristics, including student gender balance, ability status, language status, free/reduced lunch status, teacher experience levels, the number of school counselors, school grade levels (i.e. elementary, middle, high), and locale (i.e. city, suburban, town, rural). This work was intended to contribute to a growing body of research that has attempted to examine how racial/ethnic composition of place affects racially/ethnically disparate school discipline outcomes (Rocque & Paternoster, 2011; Skiba et al., 2014; Welch & Payne, 2010). It was also intended to respond to calls for evidence regarding where to intervene in inequitable and racialized school discipline outcomes (Skiba et al., 2014).

Bivariate regression with controls analysis was used to examine samples drawn from data covering all U.S. public schools during the 2013-2014 school year. To address the research question and purpose in a comprehensive manner, separate samples and regression models were constructed for all U.S. Census Bureau designated racial/ethnic groups: American Indian/Alaska Native, Asian, Black, Latinx, Native Hawaiian/Pacific Islander, Two or more Races.

Research Question

This study asks the following research question:

What effect does the proportion of White students in a school have on school-level risk of out of school suspensions for students of color as compared to White students?

This research question allowed the analysis to determine if the proportion of White students in a school is a significant predictor of the school-level discipline gap between Whites and students of color. It also allowed for the examination of how the control variables affected the relationship between school's racial/ethnic balance and racial/ethnic discipline disparities. In addition, analysis enabled a determination of the unique contribution of the predictor variables to the outcome variable while controlling for the effect of the other predictors.

Hypotheses

The primary null hypothesis in the regression analysis was that there was no relationship between the dependent variable (RRR) and the independent variables, including those designated as control variables. During the regression procedure, separate null hypotheses for each of the independent variables were tested. These tests indicated whether inclusion of each predictor improved the predicative power of the model any more than would be expected by chance.

In general, the primary alternative hypothesis in regression is that at least one of the independent variables would be useful in predicting RRR. Given this study's positioning of the proportion of White students as the focal independent variable, it was hypothesized (H_1) that the proportion of White students in a school would be a significant

predictor of the risk of suspension for all six racial/ethnic groups as compared to White students in the same school.

Data

Discipline data for this study was collected by the United States Department of Education as part of the Civil Rights Data Collection (CRDC) for the year 2013-2014. The CRDC requires biannual survey reporting from "all public schools and school districts in the United States" (U. S. Department of Education, 2016, p. 1) on key educational and civil rights issues, including school discipline. Data was obtained by written request submitted to the Office of Civil Rights in the U. S. Department of Education. This study analyzes CRDC's *school level* metrics of student disciplinary outcomes. No data on individual students, teachers, administrators, or other school-based actors was obtained or analyzed. CRDC packages discipline data with metrics related to student demographics, school size, student disability population(s), student poverty levels, and the characteristics of teachers and school professionals. CRDC obtains non-discipline data from the National Center for Education Statistics Common Core of Data. The samples used for this study were drawn from data covering all P-12 public schools in the United States *(N*=95,508) during the 2013-2014 school year.

Variables

Dependent Variable

The dependent variable in this study was a comparison of suspension rates between Whites and students of color referred to as a *relative risk ratio* (RRR) (Coutinho & Oswald, 1998; Hosp & Reschly, 2003; Parrish, 2002). RRRs compare the rate at which a condition (i.e. suspension) is applied to different sub-groups (i.e. racial/ethnic student

groups) to generate a ratio describing the extent of disparity. The RRR is generated by comparing the *risk index* (RI) (Hosp & Reschly, 2003; Reschly, 1997) for one group to the RI for another group. In this study, the RI was obtained by dividing the number of instances of suspensions for each racial/ethnic group (including White students) by the total number of students from that same group enrolled in a school (see Figure 1).

$$RRR = \frac{\frac{\text{\# of Suspension for Blacks in a School}}{\text{\# of Blacks Enrolled in a School}}}{\frac{\text{\# of Suspension for Whites in a School}}{\text{\# of Whites Enrolled in a School}}}$$

Figure 1. Relative Risk Ratio (RRR) Calculation for Black Students.

Relative risk ratio calculations yielded a series of risk indices, representing the number of suspensions per 100 students for a given racial/ethnic group. Dividing the RI of one group (i.e. Black students) by the RI of another (i.e. White students) produces an RRR, the dependent variable, which compares a group's risk of suspension to the risk of a comparison group (Donovan & Cross, 2002; McLoughlin & Noltemeyer, 2010). For instance, if 2.64% of suspensions were given to Black students, and 1.18% of suspensions were given to White students, then the RRR for Black students would be 2.24, meaning that Black students had 2.24 times the risk of suspension as White students.

In this study, White students were selected as the comparison group because, while non-White students became a majority of U.S. public school students in 2014 (U.S. Department of Education, 2015), data from the current analysis indicated that they remain

the majority in most U.S. schools. Therefore, an RRR of 1.0 indicated that the risk of suspension was the same for Whites as it was for the relevant the students of color being modeled (i.e. American Indian/Alaska Native, Asian, Black, Latinx, Native Hawaiian/Pacific Islander, Two or more Races). An RRR above 1.0 indicated that students of color were at greater risk of suspension than Whites, and an RRR less than 1.0 meant that students of color were underrepresented in suspension as compared to Whites within the same school. Previous studies have made extensive use of the RRR to assess levels of disproportionality in school discipline, primarily as a means of showing which types of students within a bounded system (school, district, etc.) had the highest risk of school discipline (American Psychological Association Zero Tolerance Task Force, 2008; Blake, Butler, Lewis & Darensbourg, 2011; Lewis, Butler, Bonner & Joubert, 2010; McLoughlin & Noltemeyer, 2010). The current study, by using RRR as the dependent variable in a series of regression models, allowed one to assess what kinds of school-level traits predict disparities between Whites and students of color across a large sample of U.S. public schools.

Focal Independent Variable

The focal independent variable in this study was the proportion of White students in a school (proportion White). Proportion White was calculated using data on school's total enrollment and the number of White students reported for each school. It was entered into the regression models as proportions, where .75 indicates that 75% of a school population was identified as White.

Previous studies have shown that a school's demographic composition affects discipline outcomes (Raffaele Mendez et al., 2002; Welch & Payne, 2010). This past

literature has focused on the proportion of Black students enrolled, finding that % Black increases the likelihood and harshness of exclusionary discipline within a school. Previous studies have also suggested that more highly-resourced, predominantly-White school contexts may produce higher rates of suspension for students of color as compared to Whites (Rausch & Skiba, 2006; Rocque & Paternoster, 2011; Wallace et al., 2008). In the only study to directly test how demographic composition affects the magnitude of disparity, Eitle & Eitle (2004) found that the size of the Black population was *not* related to the overrepresentation of Black students in suspensions.

Control Variables

Proportion of female students. The proportion of female students was calculated using data on school's total enrollment and the number of female students reported for each school. It was entered into the regression models as a decimal.

Gender has repeatedly been shown to influence suspension, with males typically receiving suspension at a higher rate (Fabelo et al., 2011; Finn & Servoss, 2013; Rocque & Paternoster, 2011). Males have also been found to misbehave more often than females (Finn et al., 2008; Skiba et al., 2002). However, no known studies have examined the effect of gender and gender balance on the racial/ethnic discipline gap within schools. Thus, gender is included in this study as a means of examining if and how gender balance influences the magnitude of the racial/ethnic discipline gap. Gender is also modeled as a control variable to provide proper estimates of disparity in schools with disproportional gender representation.

Proportion of students identified as disabled. This variable includes students identified under the Individuals with Disabilities Education Act (IDEA) (20 U.S.C. §

1412) and Section 504 of the Rehabilitation Act of 1973 [29 U.S.C. § 794 (504)]. The proportion of students identified as disabled was calculated using data on school's total enrollment and the number of students reported as having a disability under IDEA and Section 504 of the Rehabilitation Act for each school. It was entered into the regression models as a decimal.

Disability status has been shown to affect suspension rates and racial/ethnic gaps in discipline outcomes, with student disability being positively related to increased risk of suspension (Krezmien et al., 2006; Skiba, et al., 2002; Wallace, et al., 2008). Combining this documented relationship between ability status and exclusionary discipline with a separate, substantial line of research indicating that children of color tend to be over-identified in special education services (Artiles, Harry, Reschly, & Chinn, 2002; Artiles, Kozleski, Trent, Osher, & Ortiz, 2010), highlights the complex intersectionality of ability status, race, and school discipline. In this study, proportion disability is included as a control variable to account for the potential influence of variance in disability status on the dependent variable. It will also allow for an examination of whether racial/ethnic composition of place (proportion White) is related to proportion disability, and how that may affect the racial/ethnic discipline gap.

Proportion of students identified as limited English proficient. The proportion of limited English proficient (LEP) students, now more commonly known as known as English Learner students or English Language Learners, was calculated using data on school's total enrollment and the number of students identified as LEP reported for each school. It was entered into the regression models as a decimal.

There is very little research evaluating the role of language status in school discipline. While a previous found that LEP students were under-represented in the incidence of exclusionary discipline (Anderson & Ritter, 2017), language is an important component of racial/ethnic identity, and has a long history of contributing to the hegemony of the English language in U.S. schools (Heller, 2007; Kluckhohn, 1962; Wiley & Lukes, 1996). Thus, proportion LEP is of interest given the theoretical orientation of the study, and the potential for language status to emerge as a factor that influences racial/ethnic disparities in school discipline, particularly within predominantly-White school contexts.

Proportion of novice teachers. Novice teachers are here defined as those in their first or second year of teaching. The proportion of novice teachers was calculated from counts of first and second year teachers and the total number of teachers in a school. It was entered into the regression models as a decimal.

Teachers play a critical role in school discipline. Indeed, most discipline referrals originate with classroom teachers (Skiba et al., 2002), and the characteristics of teachers and their dispositions toward students have been a subject of interest for discipline scholars (Gregory & Thompson, 2010; Payne & Welch, 2010; Rocque & Paternoster, 2011). Several studies have found that more experienced teachers produce higher suspension levels for students of color after statistically controlling for various school-level variables such as racial/ethnic balance, poverty measures, student homogeneity (Eitle & Eitle, 2004; Raffaele Mendez, et al., 2002). Given that the experience levels of school teaching staffs vary across locations (U.S. Department of Education, 2009b), this

study is interested in whether proportion novice teachers as a proxy for staff experience levels can predict racial/ethnic differences in school discipline.

Full-time counselors. This variable is the number of full-time counselors employed at a school. Some schools report partial counselors because school support staff are often shared among several schools.

Past research has shown that African American students are more likely to be referred to the school counselor for behavioral concerns (Adams, Benshoff, & Harrington 2007), and that referrals to school counselors for disruptive behavior are characterized by racial/ethnic disproportionality as well (Bryan, Day-Vines, Griffin, & Moore-Thomas, 2012). While counselors do not typically carry out school discipline, they occupy a prime position to influence discipline outcomes (Grothaus, 2013). Further, they may represent an alternative to disciplinary referrals and/or an additional form of support and advocacy for students involved in disciplinary incidents (Day-Vines et al., 2012). This study is interested in whether the number of counselors in a school may serve as a protective factor against racial/ethnic disparities. In addition, given research suggesting that more highly resourced school contexts ration student support services based on student race/ethnicity (Welch & Payne, 2010), the number of school counselors may interact with White in ways that affect racial/ethnic rates of discipline.

Proportion of students eligible for free/reduced lunch. Students eligible for both free and reduced lunch were totaled and compared to the total enrollment in a school. Proportion FRL was entered into the regression models as a decimal.

Proportion FRL is intended to be a research proxy for factors related to aggregate school poverty, economic (dis)advantage, and/or socioeconomic status (SES) of the

students enrolled in a school. While there is some discussion about the use of FRL as a proxy for poverty (Sparks, 2014), the correlation between FRL and other poverty metrics is widely acknowledged (Snyder & Musu-Gillette, 2015), and many of the most methodologically rigorous school discipline studies in recent years continue to use FRL in this manner (see Fabelo et al., 2011; Skiba et al., 2014).

SES is perhaps the most well researched predictor of school discipline after student race. Indeed, the relationship of race and SES to discipline outcomes has have been extensively researched. Decades of research have found that poverty has also been found to be a consistent predictor of school discipline, with low socioeconomic status (SES) students typically receiving suspension at a higher rate (Brantlinger, 1991; Nichols, 2004; Wu et al., 1982). While SES is a consistent, strong predictor of the likelihood of suspension, and even though students of color are over-represented in poverty metrics, SES has demonstrated only a limited ability to explain racial/ethnic differences in discipline outcomes (Christle et al., 2004; Fabelo et al., 2011; Wallace et al., 2008).

Given its prominence in the literature and its covariance with school demographic composition, SES was included in this study primarily as a control variable, as a means of isolating the effect of proportion White. SES can also function as a well-known, consistent, and relatively strong predictor of discipline outcomes, SES could also function as a yardstick of sorts, against which to compare the focal independent variable.

School level. The data included four categories for the school level variable: elementary, middle, high, and other. The other category consisted of schools with grade spans such as K-8, or 6-12, that fall outside of the traditional definitions of school level.

As a categorical variable, school level was dummy coded, allowing it to be entered into the regression equation. Elementary schools were chosen as the reference category because they represent the largest proportion of U.S. public schools. Regression coefficients based on categorical variables coded in this fashion are read as compared the reference group. For instance, a coefficient for high schools of -.25 predicts that high schools would tend to have lower values than elementary schools on the dependent variable.

Research has shown that school level affects discipline outcomes. While there tends to be a higher incidence of discipline in secondary schools (Raffaele Mendez et al., 2002), research has raised concerns about discipline outcomes on the elementary level. There is evidence that elementary students may be at increased risk of exclusionary discipline (Butler, 2011), racially/ethnically motivated discipline (Rocque & Paternoster, 2011), and that patterns of overrepresentation for student of color begin at the elementary level and continue to high school (Brooks, Schiraldi, & Zeidenberg, 2001). Therefore, school level is of interest here because part of the purpose of this study is to provide some measure of prediction about where racial/ethnic discipline gaps may be largest, and where best to target equity-seeking interventions.

School locale. The data included 12 categories for the school locale variable. The National Center for Educational Statistics locale code system classifies territory into four major types: city, suburban, town, and rural. Each type has three subcategories. The city and suburb types are further divided into large, midsize, and small. Towns and rural areas are further divided into fringe, distant, or remote (U.S. Department of Education, n.d.). To avoid an excessive number of locale-related predictors, the subcategories were

collapsed into the four basic types (city, suburban, town, rural) and dummy coded. Suburban was chosen as the reference category because it represents the largest proportion of U.S. public schools. Regression coefficients based on categorical variables coded in this fashion are read as compared the reference group. For instance, a coefficient for city schools of .25 predicts that city schools would tend to have higher values than suburban schools on the dependent variable.

Research has shown that urban schools serve higher percentages of students of color (U.S. Department of Education, 2010), and that urbanicity influences school discipline, primarily because large, urban districts tend to have higher incidence of suspension (Skiba et al., 2002; Wallace et al., 2008; Welch & Payne, 2010). However, no known studies have examined how locale influences the magnitude of racial/ethnic discipline disparities. It is worth noting again that populations of color are growing in suburban, town and rural schools (U.S. Department of Education, 2007; U.S. Department of Education, 2010), and studies have suggested that growing populations of color will result in more punitive, racially/ethnically-motivated discipline based on perceptions of racial threat (Welch & Payne, 2010). Given the lack of research on race/ethnicity and discipline outcomes in non-urban, predominantly-White contexts, this study is interested in examining how locale affects the racial/ethnic discipline gap.

Model Fitting

Meaningful and Stable Estimates of the Dependent Variable

To obtain a meaningful value for the RRR, schools must have values other than 0 for the 1) students of color of the group being modeled, 2) the reference group (White students), or 3) incidence of suspension for the students of color being modeled (Bollmer

Bethel, Garrison-Mogren, & Brauen, 2007). Furthermore, estimates of disproportional representation, such as RRRs, often become unstable in the case of small samples, small student populations, or limited instances of the condition under examination (i.e. suspension) (Bollmer et al., 2007; Hosp & Reschly, 2004). While RRRs obtained from small numbers are accurate and true to the reported data, they can create extreme leverage points in regression models. To the extent that schools with extremely small numbers of students from the racial/ethnic group being modeled or the comparison group are not representative of the overall sample, these cases can be misleading in terms of generalizability of regression results. To ensure meaningful and stable values for school RRRs, criteria were established upon which the decision to retain/exclude cases was based. These criteria, and information on the retention/exclusion of cases is discussed in greater detail below in the Sample section.

Normality of the Dependent Variable

The RRR distributions for all six racial/ethnic group models were strongly positively skewed and leptokurtic. Distributions of this nature were expected for two reasons. First, as discussed above, virtually all past research has found that Black and American Indian students exhibit higher risk of suspension than White students. There is also evidence that Latinx students are at greater risk of suspension than Whites in many, but not all, studies. Previous research has not modeled RRRs for Native Hawaiian/Pacific Islander or students of two or more races, and Asians have been found to have lower risk than their White counterparts. Therefore, except for Asians, where previous scholarship is available, students of color would be expected to have RRR distributions that cluster above 1.0. Second, the nature of the RRR metric confines cases of under-selection for

students of color between 0 and 1, while cases of over-selection cluster can theoretically go out to infinity. As a result, one would expect fewer cases to fall between the values 0 and 1(indicating lower risk of suspension for White students), and most cases to begin to cluster above 1.0, rising to the mean point, and then tailing out to the right hand, positive side of the distribution peak. This clustering around the central tendency, with a longer/fatter tail on the right side very accurately describes a positively skewed and leptokurtic distribution observed in the data for this study.

In cases of non-normal variables, statistics experts recommend log transformation as a means of improving prediction and reducing the impact of outliers (Tabachnick & Fidell, 2013). As has been done in previous discipline research using risk ratios as a dependent variable (Eitle & Eitle, 2004, Hosp & Reschly, 2003), a log transformation (lg10) of the RRR was performed to approximate a normal distribution.

Sample & Selection

A separate sample was selected for each of the six racial/ethnic groups based on the criteria below. To remain in the analysis, cases had to have:

1. At least 10 students of color (of the group being modeled)
2. At least 10 White students
3. At least 1 case of out of school suspension for students of color of the group being modeled
4. At least 1 case of out of school suspension for White students

Criterion 1. A non-zero value for the enrollment of students of color (of the group being modeled) is required for a numerically meaningful RRR. Further, small values for the student groups of color being modeled tends to produce unstable estimates of RRR

because minor variations can produce drastic changes in the size of the risk ratio (Bollmer et al., 2007). Scholars have indicated that a minimum of 10 students are required for stable estimates (Bollmer et al., 2007; Bollmer, Bethel, Munk, & Bitterman, 2014). Cases with less than 10 students of color from the group being modeled were excluded from the analysis.

Criterion 2. As was the case with criterion 1, a non-zero value for White enrollment is required to for a numerically meaningful RRR. Again, small values for White students tends to produce unstable estimates of RRR (Bollmer et al., 2007). Bollmer and colleagues (2007, 2014) indicate that a minimum of 10 students are required for stable estimates. However, rather than deletion, Bollmer and colleagues (2007, 2014) suggested the calculation of an *alternate risk ratio* (ARR). ARR uses mean replacement for the comparison group to obtain meaningful risk ratios. For instance, Bollmer and colleagues (2007) examined the relative risk of receiving special education services in a sample of school districts across several states. In districts with less than 10 White students, they calculated ARR by substituting the state mean risk ratio. The current study might have used national, state, or local educational agency (LEA) means for schools with less than 10 White students. The primary benefit of using ARR is the retention of a larger sample, while the primary risk revolves around undesirable and/or unexpected changes to the variance in the dependent variable.

The use of ARR was undesirable in this study for several reasons. First, given the number of schools in the raw data for this study (~90,000) adequate sample size was not a critical concern for this analysis. Therefore, maintaining representative variation in the dependent variable *within each racial/ethnic group sample* was deemed to be of greater

importance. Secondly, a substantial proportion (between 3 and 17%) of the potential cases in each racial/ethnic group sample exhibited less than 10 White students. Here again, imputation of ARR would likely create dramatic changes in the variability of the dependent variable. Furthermore, it is unclear from theory and research whether geographic/political boundaries (state, district, LEA) have any relationship with racial/ethnic variation in school discipline outcomes. For these reasons, cases with fewer than 10 students from the comparison group (Whites), were excluded from the analysis.

Criterion 3. Unlike the other three criteria, a 0 value for suspensions of students of color does not automatically produce a meaningless value for RRR. If the other three criteria are met, a 0 value for OSS for the students of color being modeled would yield an RRR of 0, indicating a dramatic under-selection of students of color for suspension. Nonetheless, it was necessary to exclude cases with no OSS for students of color for several reasons. First, the requirement in criterion 4 that cases with no White OSS be excluded, would have the effect of biasing RRR estimates higher because schools with 0 incidence OSS for White would likely have some of the highest RRRs and some of the most racially/ethnically disparate (although numerically meaningless) discipline rates. To retain balance in RRR estimates, it was necessary to exclude schools with no OSS for students of color. Not doing so would have left a large number of schools that were extremely low suspending (see criterion 4) in the sample, while removing a substantial portion of schools that were high suspending. Therefore, I excluded schools with 0 OSS for the students of color being modeled.

Criterion 4. As with the first two criteria, a non-zero value for the condition under analysis (OSS) for White students group is required for a numerically meaningful RRR.

Similar to criterion 2, past literature has suggested using ARR for cases in which the comparison group has no incidence of the condition under analysis (OSS) (Bollmer et al., 2007; Bollmer et al., 2014). However, once again, imputation of ARR was undesirable for several reasons. On a logical level, it seems undesirable to attempt to compare the risk of discipline when the conditions underlying that risk (suspension) are not present. This study is based on school level data. Thus, in the absence of school level risk of suspension, and considering the ample sample sizes present, along with previously mentioned concerns about biasing the dependent variable, cases with no incidence of White suspension were excluded from the analysis.

Table 1 shows the starting sample, the number of cases excluded for reasons related to stability of the dependent variable, missing values, and outlier analysis, as well as the final sample for each racial/ethnic group. Additional details about the samples are presented in Appendices A and D. Sample details, comparing the mean, range, minimum, and maximum for White students and student groups of color are presented in Appendix A: Sample Details by Racial/Ethnic Group. Appendix D: The Proportion of White Students by Racial/Ethnic Group Sample shows the the mean, median and percentile values for the proportion of White students in the full sample (~90,000 schools) and each of the racial/ethnic group samples. Values for the racial/ethnic group samples are similar to those of the full sample, with the exception of NH/PI, who appear to attend schools with significantly lower proportions of White students when compared to the full sample and to students from the other racial/ethnic groups.

Data Screening

Data screening began with the identification and treatment of cases with missing or impossible values. No cases had missing data for the dependent variable, or for the dummy-coded categorical variables for school level and locale. Several continuous

Table 1

Starting Sample, Excluded Cases by Reason, and Final Sample

Starting Sample	89467					
	AI/AN	Asian	Black	Latinx	NH/PI	TomR
Stable estimates of RRR						
<10 Students of Color	84265	58422	38520	24709	90349	48226
<10 White Students	555	788	5387	5282	137	560
No Suspensions for Students of Color	4421	25809	16015	29340	2139	25746
No Suspensions White	206	633	4170	4451	101	1000
Missing/Impossible Values						
Proportion Disability	47	21	185	177	2	48
Proportion LEP	0	2	6	5	0	2
Proportion Novice Teachers	4	7	88	106	0	31
Full Time Counselors	14	0	1	1	0	0
Proportion FRL	23	75	278	271	6	115
Outlier Analysis						
+/- 3.00 standard Residual[a]	19	45	195	175	10	98
Final Sample	**4017**	**7786**	**29196**	**29526**	**1089**	**17853**

Note: LEP = Limited English Proficient, FRL = eligible for free or reduced lunch, AI/AN = American Indian/Alaska Native, NH/PI = Native Hawaiian/ Pacific/Islander, TomR = Two or more races, RRR = relative risk ratio. Cases were excluded in the order presented in the table. Figures may not represent the total number of cases in which the condition applied.

[a] The procedure for identifying these cases is explained in Ch. 4.

independent variables had cases with impossible values, such as negative proportions, or proportions that exceeded 100. Large data sets often contain impossible values because of data entry errors or variation in how those reporting data indicate the absence of value. For instance, of the schools reporting no novice teachers in 2013-2014, some schools may have reported "0," some may have entered nothing, and others may have entered an

impossible value, such as "-99," presumably as an indicator of the absence of value. During data screening, impossible values were converted to missing values.

Table 1 shows the variables the had missing or impossible values, and the number of cases excluded upon that basis for each of the six regression models. While there is not a universally accepted threshold for missing data, 5% is the most widely used point at which missing cases may be deleted with a minimal chance of biasing estimates (Garson, 2015). Missing values did not exceed 2% in any of the current regression models. Nonetheless, listwise deletion can bias data. Therefore, SPSS Missing Values Analysis was used to explore the nature of the missing data. Little's MCAR test was significant ($p<.001$) in all cases, indicating that missing data was not missing completely at random. If data are not MCAR, the usual recommendation is to avoid listwise deletion. However, experts suggest deletion *is* a reasonable (even preferable) approach in the context of regression analysis when the missing data is on predictor variables, has little relationship with the dependent variable, and the remaining sample provides sufficient power to observe hypothesized effects (Allison, 2001; Allison, 2014). These conditions applied in this study. As is evident in Table 1, all missing data was on predictor variables. Separate variance t-tests revealed that the missing data points had little relationship with the dependent variable. Sample sizes were ample for all models based on accepted thresholds (Tabachnick & Fidell, 2013). Thus, in an effort to retain 'honest' standard errors that reflect the actual amount of information used" (Allison, 2014, para. 17), cases with missing data were removed from the analysis.

Analysis

A bivariate regression with controls was conducted to assess the extent to which school-level RRR could be predicted by the proportion of white students in schools, while controlling for (a) school gender balance, (b) the proportion of students with disabilities at school, (c) the proportion of English language learners, (d) the proportion of students eligible for free or reduced lunch, (e) the proportion of novice teachers, (g) the number of school counselors, (h) school configuration (elementary, middle, high), and (i) school locale (urban, suburban, rural, town). Regression allows for the examination of how multiple predictors are jointly related to racial/ethnic gaps in suspension. It also enables predictions about how each factor (a – i above) will influence school RRR holding everything else constant. Six regression models were tested, one for each racial/ethnic group in the study. Whites were the comparison group. The regression equations for the models followed he form below:

$$\log(Y_i) = \alpha_i + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_{13} X_{13} + \varepsilon$$

$\log(Y_i)$ = Log transformed RRR (the relative risk of suspension for the students of color being modeled as compared to White students in the same school)
α_i = Constant
X_1 = proportion of White students
X_2 = proportion of females
X_3 = proportion of students identified as disabled
X_4 = proportion of students identified as limited English proficient
X_5 = proportion of novice teachers
X_6 = number of full time counselors
X_7 = proportion of students eligible for free/reduced lunch
X_8 = middle school level
X_9 = high school level
X_{10} = other school level
X_{11} = city locales
X_{12} = town locales
X_{13} = rural locales
ε_i = error/residual

While the primary focus of this analysis was on the main effect of the proportion of White students on the relative risk of suspension for student groups of color (as compared to White students), in order to isolate the effect of proportion White to the greatest extent possible, the analysis included control variables previously linked to differential outcomes in suspension rates. Likewise, the findings and interpretations discussed in subsequent chapters also emphasize the main effect of the proportion of White students on the relative risk of suspension for student groups of color (as compared to White students). The treatment of the effect of control variables within the model was limited to cases where findings diverge significantly from previous literature, or when effects related to the control variables support new or emerging findings within the school discipline literature. As such, a focus on the main effect precluded a full analysis of control variables. For instance, the addition of interaction terms to the regression models was reserved for future analysis.

Assumptions

Multivariate regression has several assumptions related to 1) sample size, 2) normality of the dependent variable and residual errors, 3) univariate and multivariate outliers, 4) linearity, 5) multicollinearity, 6) homoscedasticity, and 7) Independence of Observations/Residuals/Errors. Sample size adequacy, normality of the dependent variable, and univariate outliers were checked before the regression procedure. The remaining assumptions were tested during the procedure. Unless stated otherwise, the narrative to follow applies to all 6 racial/ethnic group models.

Sample size. Tabachnick and Fidell (2013) recommend between 20 and 40 cases for each independent variable depending on a) the type of multiple regression (standard,

step-wise, hierarchical), and b) the normality of the dependent variable (non-normal DVs require more cases). The samples in this study range between roughly 80 (NH-PI) and 2000 (Latinx) cases per independent variable.

Normality of the dependent variable (RRR). As discussed previously in this chapter, the dependent variable (RRR) differed substantially from normal in all the study's samples, which was expected based on past literature and the nature of the metric. As a result, the dependent variable was log transformed to address the issue of non-normality, as has been done in previous studies using RRR as a dependent variable (Eitle & Eitle, 2004; Hosp & Reschly, 2003) After transformation, visual inspection of a histogram approximated a normal distribution, and values for skewness and kurtosis were within accepted limits (-2 to 2) (Trochim & Donnelly, 2006).

Normal distribution of residual errors. This assumption is related to whether the *theoretical* residuals are normally distributed. Since the theoretical residuals are not available for inspection, a normal P-P plot is used to examine standardized residual values. Visual inspection found the points following the prediction line, therefore no residuals were assumed to be normally distributed.

Univariate outliers. As one would expect with large samples**,** univariate outliers were evident in the distributions for the dependent variable and several continuous independent variables, including proportion female, proportion disability, proportion limited English proficient, proportion novice teachers, and full time counselors. Notably, no significant univariate outliers were present for proportion White (the focal IV) or proportion eligible for free/reduced lunch. This demonstrates that the samples contained

relatively balanced numbers of schools across the range of values (0-100) for these metrics.

While univariate outliers can have a negative impact on regression results, there is no agreed upon method or theory for dealing with outliers (Cousineau & Chartier, 2010). However, experts have noted that regression models are rather robust to the impact of univariate outliers if analysis is based on adequate sample sizes. Therefore, the ample samples available for this study were expected to mitigate the influence of any univariate outliers (Cousineau & Chartier, 2010, Tabachnick & Fidell, 2013). This, along with the steps (described above) taken to assure stable and representative estimates of the dependent variable, suggested that while statistically "significant" outliers were present, which is almost guaranteed in samples this large, they were not expected to be particularly influential in terms of the predictive power of the model. Thus, no univariate outliers were treated during preliminary analysis.

Multivariate outliers. The influence of potential multivariate outliers was tested using 1) a χ^2 distribution of the Mahalanobis distance ($p < .001$), 2) Cook's distance (> 1.00), and 3) a standard residual score (+/- 3.00). Trial runs indicated that potential multivariate outliers identified by Mahalanobis distance had only trivial influence on results. Given their lack of influence, these cases were left in the analysis. No values across any of the six samples had a Cook's distance value over .02. However, trimming the samples at standard residual score over +/- 3.00 improved model fit (adjusted R^2) by values ranging from .005 to .014. As a result, cases with a standard residual score above 3.00 or below -3.00 were removed from the samples. When cases are removed in this manner, the residual values for the remaining scores changes accordingly, often resulting

in additional cases above a given threshold. However, only a single round of exclusions was carried out, as additional removal of cases had only trivial effects on results. Table 1 shows the number of cases removed based on extreme standard residual values.

Multicollinearity. The assumption of multicollinearity assumes that the independent variables are not highly correlated with each other. It was assessed using bivariate correlations, often with a value of .7 as a cutoff, as well as a measurement of the Variance Inflation Factor (VIF). No bivariate correlations in any of the models had a value of .7 or higher. VIF values ranged from 1.04 to 3.05 across the six models. Proportion White, proportion FRL and high school level had the highest VIF values across all models. While there is no agreed upon rules for VIF, suggested critical values for VIF range from 2.5 – 10, with the "rule of 10" being the most widely used value (O'Brien, 2007). Based on those values, no serious violations of multicollinearity were assumed.

Linearity. This assumption demands a linear relationship between the outcome variable and the independent variables. Linearity was tested using a plot of residuals against predicted values and a normal P-P plot of standardized residuals. A visual inspection of the former revealed a rectangular shape and no values outside of 3 and -3 on x and y axis. Inspection of the P-P plots found the points following the line. Based on these results, the assumption of linearity was assumed to be met.

Homoscedasticity. This assumption requires that the variance of error terms is similar across the independent variables. Homoscedasticity was tested using a plot of residuals against predicted values and a normal P-P plot of standardized residuals. A

visual inspection of the former revealed a rectangular shape and no values outside of 3 and -3 on x and y axis. Thus, no violation of homoscedasticity was recorded.

Independence of observations/residuals. Assessing the validity of the statistical independence of the observations depends mainly on how the data were collected, and how the samples were selected. The data collection factors do not apply to the current study, in which time-series data are not used. However, in the absence of random sampling, this assumption tests whether the sampling procedures contained bias. Once again, this assumption is tested using a plot of residuals against predicted values. A visual inspection of the former revealed a rectangular shape and no values outside of 3 and -3 on x and y axis, indicating that the assumption of independence was met.

CHAPTER 4: RESULTS

The primary purpose of this study was to examine the relationship between the proportion of White students in a school and school-level racial/ethnic disparities in suspension between White students and students of color. This chapter presents the findings from the quantitative analysis. First, I present descriptive findings and notable differences between schools on variables associated with school discipline outcomes and demographic composition. Next, I describe the results of the regressions that were conducted to answer the research question for each of the six racial/ethnic groups in the study individually. Finally, I present notable differences between and amongst the racial/ethnic groups.

Descriptive Findings

This section presents the descriptive findings of the study. Each of the six models is presented separately, followed by the presentation of some notable findings based on examination of similarities and differences between and amongst the models. Recall that it was necessary to log transform the dependent variable (RRR) to meet the assumptions of regression analysis. This transformation makes interpretation of the results more difficult in terms of relating the substantive meaning of descriptive findings. For instance, suppose a school had an RRR of 2.5 for Latinx students. One might find that the risk of suspension for Asians was 2.5 times higher than Whites in a school. The logRRR of the same school would be .4, which is meaningless without reference to the original unit of

measurement. While back-transformation of logged variables is possible, bias occurs when back-transforming fitted means to the original scale (Rothery, 1988). Therefore, the descriptive statistics presented below are for *non-transformed* data. However, the same sample was used for descriptive results (using non-transformed data), and inferential results (using transformed data) in the Regression section below.

American Indian/Alaska Native

The sample for American Indian/Alaska Native (AI/AN) totaled 4017 schools. Descriptive statistics, including means, standard deviations, minimums, maximums and correlations between independent variables are summarized in Table 2. Mean RRR for AI/ANs was 3.66, indicating that on average in this sample, AI/ANs had 3.66 times the risk of suspension as White students within the same school. On average, 50% of students in the schools sampled were White and just under half were female. An average of 15% of students in the schools included in the sample were considered disabled, and an average of 7% were limited English proficient. The sample mean for the proportion of novice teachers in a school was 12%, and sampled schools averaged just over two full time counselors. The mean percentage of students eligible for free or reduced lunch was just under 60%. About a quarter (25%) of the schools were middle schools, 43% were high schools, under two percent (2%) were other, and the remainder (~30%) were elementary schools. About 31% of those sampled were city schools, 21% were town schools, 25% were rural and the remainder (~23%) were suburban.

Table 2

American Indian/Alaska Native Means, Standard Deviations, and Intercorrelations Among All Variables (n=4017)

	RRR	LogRRR	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Proportion White			1												
2. Proportion Female			0.02	1											
3. Proportion Disability			0.07	-0.11	1										
4. Proportion Limited English Proficient			-0.58	-0.03	-0.13	1									
5. Proportion Novice Teachers			-0.21	-0.03	0.01	0.19	1								
6. Full Time Counselors			-0.07	0.08	-0.10	-0.10	-0.01	1							
7. Proportion Eligible for Free/Reduced Lunch			-0.63	-0.07	0.15	0.48	0.22	-0.26	1						
8. Middle School Level			0.03	-0.01	0.12	-0.04	0.01	-0.15	0.04	1					
9. High School Level			0.10	0.04	-0.12	-0.23	-0.09	0.56	-0.34	-0.50	1				
10. Other School Level			0.07	-0.01	0.00	-0.05	-0.02	-0.04	-0.02	-0.08	-0.11	1			
11. City Schools			-0.34	0.02	-0.04	0.34	0.06	0.12	0.16	-0.02	-0.04	-0.05	1		
12. Town Schools			0.20	0.00	0.02	-0.11	-0.03	-0.15	-0.02	0.07	-0.05	-0.04	-0.35	1	
13. Rural Schools			0.19	-0.04	0.11	-0.23	-0.05	-0.20	0.02	-0.05	-0.01	0.11	-0.39	-0.30	1
Mean	3.66	0.39	0.50	0.49	0.15	0.07	0.12	2.24	0.58	0.25	0.43	0.02	0.31	0.21	0.25
Standard Deviation	4.07	0.38	0.25	0.03	0.05	0.10	0.10	2.25	0.22	0.43	0.50	0.13	0.46	0.41	0.43
Minimum	0.15	-0.83	0.01	0.32	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0
Maximum	56.79	1.75	0.96	0.64	0.41	0.69	0.68	16.80	1.00	1	1	1	1	1	1

Note: RRR = relative risk ratio, LogRRR = log transformed relative risk ratio

Asian

The sample for Asian totaled 7786 schools. Descriptive statistics, including means, standard deviations, minimums, maximums and correlations between independent variables are summarized in Table 3. Mean RRR for Asians was 1.77, indicating that on average in this sample, Asians had 1.77 times the risk of suspension as White students within the same school. On average, 44% of students in the schools sampled were White and just under half were female. An average of 14% of students in the schools included in the sample were considered disabled, and an average of 11% were limited English proficient. The sample mean for the percentage of novice teachers in a school was 11%, and sampled schools averaged almost three full time counselors. The mean percentage of students eligible for free or reduced lunch was just under 50%. About a third (33%) of the schools were middle schools, 40% were high schools, under two percent (2%) were other, and the remainder (~27%) were elementary schools. About 38% of those sampled were city schools, 5% were town schools, 6% were rural and the remainder (~51%) were suburban.

Black

The sample for Blacks totaled 29,196 schools. Descriptive statistics, including means, standard deviations, minimums, maximums and correlations between independent variables are summarized in Table 4. Mean RRR for Black students was 3.79, indicating that on average in this sample, Blacks had 3.79 times the risk of suspension as White students within the same school. On average, 47% of students in the schools sampled were White and just under half were female. An average of 15% of students in the schools included in the sample were considered disabled, and an average of 9% wer

Table 3

Asian Means, Standard Deviations, and Intercorrelations Among All Variables (N= 7786)

	RRR	LogRRR	1	2	3	4	5	6	7	8	9	10	11	12
1. Proportion White			1											
2. Proportion Female			0.10	1										
3. Proportion Disability			0.12	-0.17	1									
4. Proportion Limited English Proficient			-0.59	-0.12	-0.14	1								
5. Proportion Novice Teachers			-0.16	-0.02	0.07	0.08	1							
6. Full Time Counselors			0.14	0.08	0.09	-0.31	-0.01	1						
7. Proportion Eligible for Free/Reduced Lunch			-0.69	-0.11	0.08	0.56	0.18	-0.21	1					
8. Middle School Level			0.00	-0.02	0.15	-0.10	0.03	-0.20	0.03	1				
9. High School Level			0.17	0.07	0.00	-0.30	-0.02	0.67	-0.18	-0.57	1			
10. City Schools			-0.27	0.00	-0.02	0.22	0.06	-0.11	0.27	0.00	-0.09	1		
11. Town Schools			0.10	0.00	0.01	-0.06	-0.01	-0.04	0.04	-0.03	0.06	-0.17	1	
12. Rural Schools			0.16	0.00	-0.01	-0.12	0.00	0.04	-0.07	-0.03	0.09	-0.20	-0.06	1
Mean	1.77	0.03	0.44	0.49	0.14	0.11	0.11	2.82	0.47	0.33	0.40	0.38	0.05	0.06
Standard Deviation	2.43	0.43	0.26	0.02	0.05	0.12	0.09	2.55	0.26	0.47	0.49	0.49	0.21	0.25
Minimum	0.02	-1.61	0.01	0.34	0.00	0.00	0.00	0.00	0	0	0	0	0	0
Maximum	37.90	1.58	0.97	0.65	0.40	0.79	0.67	18.60	1	1	1	1	1	1

Note: RRR = relative risk ratio, LogRRR = log transformed relative risk ratio

limited English proficient. The sample mean for the percentage of novice teachers in a school was 12%, and sampled schools averaged just under two full time counselors. The mean percentage of students eligible for free or reduced lunch was 57%. About a quarter (26%) of the schools were middle schools, 25% were high schools, under two percent (2%) were other, and the remainder (~47%) were elementary schools. About 33% of those sampled were city schools, 11% were town schools, 16% were rural and the remainder (~40%) were suburban.

Latinx

The sample for Latinx totaled 29,526 schools. Descriptive statistics, including means, standard deviations, minimums, maximums and correlations between independent variables are summarized in Table 5. Mean RRR for Latinxs was 2.03, indicating that on average in this sample, Latinxs had 2.03 times the risk of suspension as White students within the same school. On average, 47% of students in the schools sampled were White and just under half were female. An average of 14% of students in the schools included in the sample were considered disabled, and an average of 11% were limited English proficient. The sample mean for the percentage of novice teachers in a school was 12%, and sampled schools averaged just under 2 full time counselors. The mean percentage of students eligible for free or reduced lunch was 56%. About a quarter (27%) of the schools were middle schools, 27% were high schools, one percent (1%) were other, and the remainder (~45%) were elementary schools. About 32% of those sampled were city schools, 13% were town schools, 16% were rural and the remainder (~39%) were suburban.

Table 4

Black Means, Standard Deviations, and Intercorrelations Among All Variables (N= 29,196)

	RRR	LogRRR	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Proportion White			1												
2. Proportion Female			0.03	1											
3. Proportion Disability			0.09	-0.18	1										
4. Proportion Limited English Proficient			-0.54	-0.06	-0.16	1									
5. Proportion Novice Teachers			-0.22	0.00	0.03	0.12	1								
6. Full Time Counselors			0.10	0.06	0.02	-0.19	-0.05	1							
7. Proportion Eligible for Free/Reduced Lunch			-0.63	-0.08	0.07	0.41	0.20	-0.30	1						
8. Middle School Level			0.05	0.01	0.12	-0.12	0.00	-0.01	-0.06	1					
9. High School Level			0.14	0.05	-0.02	-0.22	-0.04	0.60	-0.23	-0.34	1				
10. Other School Level			0.02	-0.01	0.02	-0.05	0.02	-0.01	0.00	-0.08	-0.08	1			
11. City Schools			-0.34	0.03	0.00	0.23	0.09	-0.07	0.24	-0.05	-0.08	0.00	1		
12. Town Schools			0.14	0.00	0.01	-0.11	-0.01	-0.05	0.07	0.03	0.04	-0.04	-0.24	1	
13. Rural Schools			0.22	0.00	-0.01	-0.18	-0.04	-0.01	-0.01	0.00	0.05	0.07	-0.31	-0.15	1
Mean	3.79	0.43	0.47	0.49	0.15	0.09	0.12	1.88	0.57	0.26	0.25	0.02	0.33	0.11	0.16
Standard Deviation	3.98	0.35	0.27	0.03	0.05	0.12	0.11	1.87	0.25	0.44	0.43	0.14	0.47	0.31	0.37
Minimum	0.15	-0.83	0.01	0.25	-0.03	0.00	0.00	0.00	0.00	0	0	0	0	0	0
Maximum	55.73	1.75	0.98	0.71	0.57	0.76	0.84	14.45	1.00	1	1	1	1	1	1

Note: RRR = relative risk ratio, LogRRR = log transformed relative risk ratio

Table 5

Latinx Means, Standard Deviations, and Intercorrelations Among All Variables (N= 29,526)

	RRR	LogRRR	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Proportion White			1												
2. Proportion Female			0.03	1											
3. Proportion Disability			0.12	-0.16	1										
4. Proportion Limited English Proficient			-0.59	-0.05	-0.19	1									
5. Proportion Novice Teachers			-0.20	0.00	0.03	0.11	1								
6. Full Time Counselors			0.09	0.06	0.05	-0.24	-0.04	1							
7. Proportion Eligible for Free/Reduced Lunch			-0.65	-0.07	0.05	0.50	0.20	-0.27	1						
8. Middle School Level			0.07	0.02	0.13	-0.15	0.01	-0.01	-0.06	1					
9. High School Level			0.16	0.02	-0.02	-0.26	-0.05	0.58	-0.24	-0.37	1				
10. Other School Level			0.03	-0.01	0.01	-0.04	0.02	-0.01	-0.02	-0.07	-0.07	1			
11. City Schools			-0.34	0.02	0.00	0.20	0.08	-0.04	0.24	-0.05	-0.09	0.01	1		
12. Town Schools			0.16	0.00	0.00	-0.08	-0.01	-0.07	0.04	0.04	0.05	-0.04	-0.26	1	
13. Rural Schools			0.25	-0.01	-0.01	-0.15	-0.03	-0.03	-0.06	-0.01	0.08	0.05	-0.30	-0.17	1
Mean	2.03	0.11	0.47	0.49	0.14	0.11	0.12	1.83	0.56	0.27	0.27	0.01	0.32	0.13	0.16
Standard Deviation	2.46	0.41	0.27	0.03	0.05	0.14	0.11	1.91	0.25	0.45	0.44	0.12	0.47	0.33	0.37
Minimum	0.02	-1.63	0.01	0.27	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0
Maximum	33.18	1.52	0.98	0.69	0.52	0.86	0.84	15.00	1.00	1	1	1	1	1	1

Note: RRR = relative risk ratio, LogRRR = log transformed relative risk ratio

Native Hawaiian/Pacific Islander

The sample for Native Hawaiian/Pacific Islander (NH/PI) totaled 1089 schools. Descriptive statistics, including means, standard deviations, minimums, maximums and correlations between independent variables are summarized in Table 6. Mean RRR for NH/PI s was 3.27, indicating that on average in this sample, NH/PIs had 3.27 times the risk of suspension as White students within the same school. On average, 31% of students in the NH/PI schools sampled were White and just under half were female. An average of 13% of students in the NH/PI were considered disabled, and an average of 15% were limited English proficient. The sample mean for the percentage of novice teachers in a school was 11%, and sampled schools averaged just under three full time counselors. The mean percentage of students eligible for free or reduced lunch was just under 60%. About 30% of the schools in the sample were middle schools, 43% were high schools, and the remainder (~27%) were elementary schools. Almost half (46%) those sampled were city schools, 8% were town schools, 5% were rural and the remainder (~41%) were suburban.

Two or More Races

The sample for Two or more Races (TomR) totaled 17,853 schools. Descriptive statistics, including means, standard deviations, minimums, maximums and correlations between independent variables are summarized in Table 7. Mean RRR for TomR was 3.83, indicating that on average in this sample, student of TomR had 3.83 times the risk of suspension as White students within the same school. On average, 52% of students in the schools sampled were White and just under half were female. An average of 15% of students in the schools included in the sample were considered disabled, and an average

Table 6

Native Hawaiian/Pacific Islander Means, Standard Deviations, and Intercorrelations Among All Variables (N= 1089)

	RRR	LogRRR	1	2	3	4	5	6	7	8	9	10	11	12
1. Proportion White			1											
2. Proportion Female			0.15	1										
3. Proportion Disability			0.10	-0.15	1									
4. Proportion Limited English Proficient			-0.53	-0.10	-0.16	1								
5. Proportion Novice Teachers			-0.14	-0.06	0.11	0.09	1							
6. Full Time Counselors			0.09	0.10	-0.06	-0.37	-0.02	1						
7. Proportion Eligible for Free/Reduced Lunch			-0.59	-0.12	0.09	0.59	0.13	-0.34	1					
8. Middle School Level			0.03	0.03	0.17	-0.10	-0.02	-0.21	0.05	1				
9. High School Level			0.16	0.07	-0.09	-0.37	-0.09	0.70	-0.35	-0.56	1			
10. City Schools			-0.10	0.03	0.05	0.16	-0.12	-0.05	0.18	-0.01	-0.06	1		
11. Town Schools			0.01	-0.05	0.02	-0.06	0.13	-0.09	0.05	-0.02	-0.04	-0.27	1	
12. Rural Schools			0.05	0.03	-0.06	-0.08	0.03	0.02	-0.04	-0.02	0.06	-0.21	-0.07	1
Mean	3.27	0.34	0.31	0.48	0.13	0.15	0.11	2.80	0.58	0.30	0.42	0.46	0.08	0.05
Standard Deviation	3.75	0.39	0.22	0.02	0.04	0.13	0.09	2.43	0.23	0.46	0.49	0.50	0.27	0.22
Minimum	0.13	-0.88	0.01	0.37	0.03	0.00	0.00	0.00	0.01	0	0	0	0	0
Maximum	45.54	1.66	0.93	0.58	0.34	0.81	0.53	14.00	1.00	1	1	0	1	1

Note: RRR = relative risk ratio, LogRRR = log transformed relative risk ratio

Table 7

Two or More Races Means, Standard Deviations, and Intercorrelations Among All Variables (N= 17853)

	RRR	LogRRR	1	2	3	4	5	6	7	8	9	10	11	12
1. Proportion White			1											
2. Proportion Female			0.03	1										
3. Proportion Disability			0.07	-0.17	1									
4. Proportion Limited English Proficient			-0.55	-0.05	-0.12	1								
5. Proportion Novice Teachers			-0.21	-0.02	0.06	0.10	1							
6. Full Time Counselors			0.03	0.07	-0.02	-0.18	-0.02	1						
7. Proportion Eligible for Free/Reduced Lunch			-0.59	-0.08	0.15	0.42	0.18	-0.28	1					
8. Middle School Level			0.04	0.01	0.12	-0.11	0.01	-0.03	-0.06	1				
9. High School Level			0.13	0.06	-0.05	-0.22	-0.05	0.64	-0.26	-0.39	1			
10. City Schools			-0.34	0.02	0.02	0.23	0.08	-0.06	0.24	-0.05	-0.08	1		
11. Town Schools			0.19	-0.02	0.03	-0.11	-0.01	-0.07	0.05	0.01	0.03	-0.26	1	
12. Rural Schools			0.23	-0.01	-0.01	-0.16	-0.05	-0.01	-0.05	0.00	0.05	-0.30	-0.16	1
Mean	3.83	0.43	0.52	0.49	0.15	0.08	0.12	2.00	0.54	0.29	0.27	0.33	0.12	0.16
Standard Deviation	4.01	0.35	0.26	0.03	0.05	0.10	0.11	1.93	0.24	0.45	0.45	0.47	0.32	0.36
Minimum	0.14	-0.84	0.01	0.33	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0
Maximum	45.10	1.65	0.98	0.68	0.48	0.68	0.75	13.00	1.00	1	1	1	1	1

Note: RRR = relative risk ratio, LogRRR = log transformed relative risk ratio

of 8% were limited English proficient. The sample mean for the percentage of novice teachers in a school was 12%, and sampled schools averaged two full time counselors. The mean percentage of students eligible for free or reduced lunch was 54%. About a quarter (29%) of the schools were middle schools, 27% were high schools and the remainder (~44%) were elementary schools. About 33% of those sampled were city schools, 12% were town schools, 16% were rural and the remainder (~39%) were suburban.

Between & Among Group Comparisons

This section examines the study variables across all six racial/ethnic groups. Table 8 presents sample sizes and mean values for dependent and independent variables by racial/ethnic group, allowing cross-group comparisons of school traits.

Table 8 shows that students of color from all racial/ethnic groups are over-selected for suspension in comparison to White students in the same schools. For the schools examined, there were 1.77 suspensions of Asians, 2.03 suspensions of Latinxs, 3.27 suspensions of NH/PIs, 3.66 suspensions of AI/ANs, 3.79 suspensions of Blacks, and 3.83 suspensions of students of two or more races for each White student in school. Schools in the Asian and NH/PI samples appear to share several traits relative to the other racial/ethnic groups, including higher mean numbers of full time counselors, fewer novice teachers, and a tendency toward suburban and city locales. These traits suggest that schools in the Asian and NH/PI schools may tend to be located in population centers. Asian schools, though not NH/PI schools, had notably lower means for proportion FRL. Schools in the Black, Latinx and TomR samples share several traits relative to the other racial/ethnic groups, including larger proportions of elementary schools and

Table 8

Mean Values for Variables by Racial/Ethnic Group

	AI/AN	Asian	Black	Latinx	NH/PI	TomR
	n=4017	*n*=7786	*n*=29196	*n*=29526	*n*=1089	*n*=17853
Dependent Variable						
RRR	3.66	1.77	3.79	2.03	3.27	3.83
logRRR	0.39	0.03	0.43	0.11	0.34	0.43
Independent Variables						
Proportion White	0.50	0.44	0.47	0.47	0.31	0.52
Proportion Female	0.49	0.49	0.49	0.49	0.48	0.49
Proportion Disability	0.15	0.14	0.15	0.14	0.13	0.15
Proportion LEP	0.07	0.11	0.09	0.11	0.15	0.08
Proportion Novice Teachers	0.12	0.11	0.12	0.12	0.11	0.12
Full Time Counselors	2.24	2.82	1.88	1.83	2.80	2.00
Proportion FRL	0.58	0.47	0.57	0.56	0.58	0.54
Elementary School Level[a]	0.30	0.27	0.47	0.45	0.28	0.44
Middle School Level	0.25	0.33	0.26	0.27	0.30	0.29
High School Level	0.43	0.40	0.25	0.27	0.42	0.27
Other School Level	0.02	-	0.02	0.01	-	-
City Schools	0.31	0.38	0.33	0.32	0.46	0.33
Suburban Schools[b]	0.23	0.51	0.40	0.39	0.41	0.39
Town Schools	0.21	0.05	0.11	0.13	0.08	0.12
Rural Schools	0.25	0.06	0.16	0.16	0.05	0.16

Note: LEP = Limted English Proficient, FRL = eligible for free or reduced lunch

[a] [b] Elementary schools and suburban schools were reference groups for categorical variable blocks.

racial/ethnic groups, including larger proportions of elementary schools and correspondingly lower proportions of middle and high schools. Black, Latinx and TomR schools also have very similar proportions across the locale designations, suggesting that students from these groups tend to have relatively similar residential patterns. The schools in the AI/AN sample did not fit particularly well with either of the groups, based

notably smaller proportions of White students, and a much stronger tendency toward town and rural locales.

Table 9 shows the RRRs of all six racial/ethnic groups as a function of predominantly-White status (>50% White student enrollment).

Table 9

School Level Relative Risk Ratio of Suspension by Predominantly-White Status

	AI/AN	Asian	Black	Latinx	NH/PI	TomR
Not Predominantly-White	2.85	1.16	2.57	1.16	2.60	2.86
Predominantly-White	4.33	2.55	5.05	2.93	5.51	4.60
Additional Risk of Suspension in Predominantly-White Schools[a]	1.52	2.20	1.96	2.53	2.12	1.61

[a]This value indicates how many times higher the relative risk of suspension was in predominantly-White schools as compared to non-predominantly-White schools.

In non-predominantly-White schools, Whites are suspended almost as often as Asians and Latinxs, while over 2.5 AI/ANs, Blacks, NH/PIs, and TomRs are suspended for each White suspension in non-predominantly-White schools. Predominantly-White status increases the racial/ethnic disparity of suspension in all racial/ethnic groups by between roughly 1.5 and 2.5 orders of magnitude. Specifically, in PW schools, an additional 1.52 (AI/AN), 1.61 (TomR), 1.96 (Black), 2.12 (NH/PI), 2.20 (Asian), and 2.53 (Latinx) students are suspended for every White suspension over and above the racial/ethnic disparities in NPW schools. It is notable that the groups with the lowest overall RRR (Asians & Latinxs) have the largest jump in racial/ethnic difference based on predominantly-White status.

Sample details, comparing the mean, range, minimum, and maximum for White students and student groups of color as a function of predominantly-White status are presented in Appendix B: Sample Size by Predominantly-White Status.

The RRRs of all six racial/ethnic groups by the proportion of White students (quartiles) are displayed in Appendix C: School Level Relative Risk Ratio of Suspension by the Proportion of White Students (Quartiles). Results show that across all racial/ethnic groups and quartiles, the relative risk of suspension of student groups of color (as compared to Whites) increases with the proportion of White students in the same schools. However, Asian and Latinx appear to be at a lower risk of suspension (as compared to Whites) in schools from the first quartile (0-25% White).

Table 10 presents the number and percentage of school above/below the 1.00 threshold for RRR. Recall that an RRR of 1.00 indicates racial/ethnic equity in suspension, positive values represent over-selection of students of color, and negative values represent under-selection of students of color.

Table 10

Percentage of Schools with Racial/Ethnic (Dis)proportionality and Relative Risk Ratios

RRR[a]	AI/AN (n = 4017)	Asian (n = 7786)	Black (n = 29196)	Latinx (n = 29526)	NH/PI (n =1089)	TomR (n = 17853)
<1.00	15%	48%	10%	37%	17%	9%
1.00	<1%	<1%	<1%	<1%	<1%	<1%
>1.00	85%	52%	90%	63%	83%	91%
Mean	3.66	1.77	3.79	2.03	3.27	3.83
Median	2.45	1.04	2.63	1.32	2.12	2.57

[a] < 1.00 indicates under-selection of students of color for suspension, 1.00 indicates racial/ethnic proportionality in suspension, and > 1.00 indicates over-selection of students of color for suspension.

Black and TomR students are over-selected for suspension in about 90% of schools in their respective samples. AI/AN and NH/PI students are over-selected in approximately 85% of schools. Latinxs are over-selected in over 60% of schools, while only just over half of the schools sampled over-select Asians for suspension. While all student groups were found to be over-selected for suspension compared to Whites in the same schools, these numbers suggest that racial/ethnic disparities in suspension are also pervasive and widespread for AI/AN, Black, NH/PI, and TomR student groups, and for Latinxs to a lesser degree. In other words, there does not appear to be a small plurality of schools in which these groups are particularly vulnerable to racial/ethnic disproportional discipline practices. Rather, over-selection of AI/AN, Black, NH.PI, and TomR student appears to be present in the vast majority (>80%) of schools, a strong majority for Latinxs (>60%), and a simple majority for Asians (>50%) Comparison with mean RRR across the full sample of each group shows that the percentage of schools that suspend students of color more often than their White counterparts corresponds to the magnitude of racial/ethnic disparity, such that a higher percentage of school that over-select students of color, the higher the mean RRR.

Regression

Bivariate regression with controls was used to test the association between the school-level relative risk of suspension for students of color as compared to White students and the proportion of White students in a school while controlling for several variables related to school discipline outcomes. Whereas the descriptive results above used standard RRR, the regression results presented below are based on a log transformed dependent variable (logRRR). This study asked the following research question:

What effect does the proportion of White students in a school have on school-level risk of out of school suspensions for students of color as compared to White students?

As was the case with the descriptive results, this section will present regression results for each of the six racial/ethnic groups in the study separately before presenting findings based on comparisons between and among racial/ethnic groups. Please note that almost all variables achieved accepted thresholds for statistical significance, but in many cases low semi-partial correlations made their substantive contribution questionable in the context of the large samples used in the analysis. Therefore, only noteworthy results for the control variables are presented.

American Indian/Alaska Native (AI/AN)

Table 11 summarizes the regression comparing the risk of suspension between AI/AN and White students in a sample of 4017 schools. The strongest predictors of logRRR in the AI/AN model were proportion FRL ($\beta = -0.292$, $p = .000$) and proportion White ($\beta = 0.280$, $p = .000$), such that higher levels of school FRL predicted a smaller racial/ethnic discipline gap, and larger proportions of Whites predicting a larger school discipline gap. All school levels (middle, high, other) were significant predictors of lower logRRR. This means that for AI/ANs, elementary schools tended to have larger racial/ethnic discipline gaps. This was especially pronounced in the comparison between elementary and high schools. City locales were not significantly different than suburban locales, but town and rural locales both predicted significantly lower racial/ethnic discipline gaps in comparison to suburban locales in the AI/AN model. The adjusted R^2 calculation indicated that the model predicted 26.5% of the variation in logRRR.

Table 11

The Impact of the Proportion of White Students in a School and Selected Control Variables on American Indian/Alaska Native Risk of Suspension

	β	*sr*$_i$	*t*-value	*p*-value
Intercept (*B*=0.583			5.396	<0.001
Proportion White	**0.280**	**0.179**	**13.254**	**<0.001**
Proportion FRL	-0.292	-0.191	-14.108	<0.001
Middle School Level	-0.084	-0.069	-5.110	<0.001
High School Level	-0.188	-0.124	-9.203	<0.001
City Schools	0.032	0.024	1.787	0.074
Town Schools	-0.133	-0.102	-7.525	<0.001
Rural Schools	-0.175	-0.128	-9.462	<0.001
Adjusted R^2 = 0.265				

Note: FRL = eligible for free or reduced lunch

Asian

Table 12 summarizes the regression comparing the risk of suspension between Asian and White students in a sample of 7786 schools. Proportion White (β = 0. 0.485, p = <.001) was by far the strongest predictor of logRRR in the Asian model, such that higher proportions of White students predicted larger disparities in suspension between Asian and White students in the same school. Middle school level (β =-0.222, p = .000) and high school level (β = -0.322, p = .000) were the next strongest predictors of logRRR. Percent LEP was also significant (β = -0.172, p = .000), predicting lower logRRR as the percentage of LEP students rises. Despite its low semi-partial correlation, proportion FRL is presented because the effect of proportion FRL in the Asian model is dramatically lower than is the case in other groups. The adjusted R^2 calculation indicated

that the model predicted 32.6% of the variation in logRRR.

Table 12

The Impact of the Proportion of White Students in a School and Selected Control Variables on Asian Risk of Suspension

	β	sr_i	*t*-value	*p*-value
Intercept (*B*=-0.476)			-5.311	<0.001
Proportion White	**0.485**	**0.313**	**33.636**	**<0.001**
Proportion LEP	-0.172	-0.121	-13.004	<0.001
Proportion FRL	0.085	0.055	5.942	<0.001
Middle School Level	-0.222	-0.164	-17.599	<0.001
High School Level	-0.322	-0.182	-19.607	<0.001
Adjusted R^2 = 0.326				

Note: LEP = Limted English Proficient, FRL = eligible for free or reduced lunch

Black

Table 13 summarizes the regression comparing the risk of suspension between Black and White students in a sample of 29,196 schools. The strongest predictors of logRRR in the Black model were proportion White (β = 0. 358, p = .000) and proportion FRL (β = -0.196, p = .000), such that higher proportions of White students predicted a larger racial/ethnic discipline gap and higher proportion FRL predicted a smaller racial/ethnic discipline gap. Once again, all school levels tended to have smaller racial/ethnic discipline gaps when compared elementary schools, although only high schools (β = -0.089, p = .000) predicted significantly and substantially lower logRRR. The adjusted R^2 calculation indicated that the model predicted 30.7% of the variation in logRRR.

Table 13

The Impact of the Proportion of White Students in a School and Selected Control Variables on Black Risk of Suspension

	β	sr_i	*t*-value	*p*-value
Intercept (*B*=0.352)			10.849	<0.001
Proportion White	**0.358**	**0.234**	**47.921**	**<0.001**
Proportion FRL	-0.287	-0.196	-40.234	<0.001
Middle School Level	-0.011	-0.009	-1.900	0.057
High School Level	-0.129	-0.089	-18.230	<0.001
Other School Level	-0.032	-0.031	-6.393	<0.001
Adjusted R^2 = 0.307				

Note: FRL = eligible for free or reduced lunch

Latinx

Table 14 summarizes the regression comparing the risk of suspension between Latinx and White students in a sample of 29,526 schools. The strongest predictors of logRRR in the Latinx model were proportion White (β = 0. 338, p = .000) and proportion LEP (β = -0.238, p = .000), such that higher proportions of White students predicted a larger racial/ethnic discipline gap and higher proportion LEP predicted a smaller racial/ethnic discipline gap. Percent FRL (β = -0.213, p = .000) was also a strong predictor of logRRR in the Latinx sample, such that higher proportion FRL predicted a smaller racial/ethnic discipline gap. The adjusted R^2 calculation indicated that the model predicted 44% of the variation in logRRR.

Table 14

The Impact of the Proportion of White Students in a School and Selected Control Variables on Latinx Risk of Suspension

	β	*sr*$_i$	*t*-value	*p*-value
Intercept (*B*=0.036)			1.054	0.292
Proportion White	**0.338**	**0.211**	**48.539**	**<0.001**
Proportion LEP	-0.238	-0.178	-40.777	<0.001
Proportion FRL	-0.213	-0.144	-32.985	<0.001
Adjusted R^2 = 0.440				

Note: LEP = Limted English Proficient, FRL = eligible for free or reduced lunch

Native Hawaiian/Pacific Islander (NH/PI)

Table 15 summarizes the regression comparing the risk of suspension between NH/PI and White students in a sample of 1089 schools. The strongest predictors of logRRR in the Native Hawaiian/Pacific Islander model were proportion White (β = 0. 244, p = .000) and proportion FRL (β = -0.199, p = .000) and, such that higher proportions of White students predicted a larger racial/ethnic discipline gap and higher proportion FRL predicted a smaller racial/ethnic discipline gap. The adjusted R^2 calculation indicated that the model predicted 29.7% of the variation in logRRR.

Table 15

The Impact of the Proportion of White Students in a School and Selected Control Variables on Native Hawaiian/Pacific Islander Risk of Suspension

	β	*sr*$_i$	*t*-value	*p*-value
Intercept (*B*=-0.328)			-1.338	0.181
Proportion White	**0.329**	**0.244**	**9.615**	**<0.001**
Proportion FRL	-0.289	-0.199	-7.831	<0.001
Adjusted R^2 = 0.297				

Note: FRL = eligible for free or reduced lunch

Two or more Races (TomR)

Table 16 summarizes the regression comparing the risk of suspension between TomR and White students in a sample of 17,853 schools. In the TomR model, the strongest predictor of school-level logRRR was high school level (β = -0. 169, p = .000), followed by proportion White (β = 0. 161, p = .000). Middle school level and Other school level were also significant predictors of lower racial/ethnic gaps in suspension between TomR and White students in the same school. As in all the previous models, higher proportion FRL (β = -0.153, p = .000) was a strong predictor of lower logRRR. The adjusted R^2 calculation indicated that the model predicted 17.6% of the variation in logRRR.

Table 16

The Impact of the Proportion of White Students in a School and Selected Control Variables on Students of Two or More Races' Risk of Suspension

	β	sr_i	*t*-value	*p*-value
Intercept (*B*=0.285)			5.722	<0.001
Proportion White	**0.161**	**0.243**	**23.761**	**<0.001**
Proportion FRL	-0.153	-0.217	-22.491	<0.001
Middle School Level	-0.130	-0.153	-19.063	<0.001
High School Level	-0.169	-0.262	-24.851	<0.001
Adjusted R^2 = 0.176				

Note: FRL = eligible for free or reduced lunch

Between & Among Group Comparisons (Regression Models)

Table 17 presents standardized regression coefficients *(β),* semi-partial correlations *(sr_i), and* R^2 calculations for the six racial/ethnic group models.

Table 17

Regression Outcomes Across Racial/Ethnic Groups

		AI/AN	Asian	Black	Latinx	NH/PI	TomR
Proportion White	β	0.280***	0.485***	0.358***	0.338***	0.329***	0.161***
	sr_i	0.179	0.313	0.234	0.211	0.244	0.243
Proportion Female	β	0.007	0.036***	0.014**	0.009	0.085**	0.031***
	sr_i	0.007	0.035	0.014	0.008	0.082	0.032
Proportion Disability	β	-0.085***	0.021*	0.012*	0.043***	0.023	0.006
	sr_i	-0.080	0.019	0.011	0.040	0.021	0.006
Proportion LEP	β	0.050**	-0.172***	0.038***	-0.238***	0.094*	0.023**
	sr_i	0.037	-0.121	0.030	-0.178	0.064	0.029
Proportion Novice Teachers	β	0.007	0.056***	-0.015**	-0.010*	-0.030	0.004
	sr_i	0.007	0.055	-0.014	-0.009	-0.028	0.004
Full Time Counselors	β	0.140***	-0.076***	0.050***	0.048***	-0.014	0.027***
	sr_i	0.105	-0.053	0.037	0.036	-0.009	0.038
Proportion FRL	β	-0.292***	0.085***	-0.287***	-0.213***	-0.289***	-0.153***
	sr_i	-0.191	0.055	-0.196	-0.144	-0.199	-0.217
Middle School Level	β	-0.084***	-0.222***	-0.011	0.005	0.006	-0.130***
	sr_i	-0.069	-0.164	-0.009	0.004	0.004	-0.153
High School Level	β	-0.188***	-0.322***	-0.129***	-0.081***	0.018	-0.169***
	sr_i	-0.124	-0.182	-0.089	-0.055	0.009	-0.262
Other School Level	β	-0.047**	-0.047***	-0.032***	-0.011*	-	-
	sr_i	-0.045	-0.043	-0.031	-0.011	-	-
City Schools	β	0.032	0.040***	0.026***	0.044***	-0.057*	-0.030***
	sr_i	0.024	0.038	0.023	0.038	-0.052	-0.034
Town Schools	β	-0.133***	0.039***	0.023***	-0.015**	-0.139***	0.019**
	sr_i	-0.102	0.037	0.020	-0.013	-0.130	0.022
Rural Schools	β	-0.175***	-0.175***	-0.040***	-0.004	-0.047	0.024***
	sr_i	-0.128	-0.128	-0.035	-0.003	-0.045	0.027
Adjusted R^2 (full models)		0.265	0.326	0.307	0.440	0.297	0.176

*p<.05, **p<.01, ***p<.001

Note: AI/AN = American Indian/Alaska Native, NH/PI = Native Hawaiian/ Pacific Islander, TomR = Two or more races, LEP = Limted English Proficient, FRL = eligible for free or reduced lunch

Proportion White. Higher proportions of White students predicted higher disparities in suspension between students of color and White students across all models. This finding provides strong empirical evidence to support the study's primary hypothesis. Proportion White also accounted for substantial variance in logRRR for students of color across all six models. It was the strongest predictor in four models (Asian, Black, Latinx, NH/PI), and the second strongest predictor in the remaining two models. Notably, proportion White has the highest semi-partial correlation (sr_i = *.313, or 31.3%)* in the Asian model, indicating that, holding all other predictors constant, proportion White accounts for more variance in logRRR for Asians than for any other racial/ethnic group in the analysis.

Proportion female. Higher proportions of female students predicted higher disparities in suspension between students of color and White students across all models. However, proportion female was not a significant predictor of logRRR in 2 (AI/AH & Latinx) of the 6 models, and a relatively weak (although significant) predictor in the remaining models.

Proportion disability. Percent disability was not a significant predictor of logRRR in 2 (NH/PI & TomR) of the six models, and a relatively weak (although significant) predictor in the remaining models. Higher proportions of disability predicted slightly higher logRRR in three models (Asian, Black, Latinx) and lower logRRR in the AI/AN model.

Proportion LEP. Percent LEP students accounted for a substantial proportion of the variance in logRRR in the Asian (sr_i = - *.121, or 12.1%)* and Latinx (sr_i = - *.178, or 17.8%)* models, both of which predicted that higher proportions of LEP students would

be associated with lower logRRR. Percent LEP was a relatively weak (although significant) predictor in the remaining models. In contrast to the Asian and Latinx models, these models predicted that more LEP students would be associated with *higher* logRRR.

Proportion novice teachers. The percentage of novice teachers was among the weakest predictors of logRRR across the six models. It was significant in only three (Asian, Black, Latinx), and only made a substantial contribution to the Asian model (sr_i = *.055, or 5.5%)*. In the Asian model, higher proportions of novice teachers predicted higher logRRR, while the Black and Latinx model predict lower logRRR.

Full time counselors. The number of full time counselors was a significant predictor of logRRR in all but one (NH/PI) of the models. However, only accounted for a substantial amount of variance in logRRR in the AI/AN model. In the significant models, additional counselors were associated with higher logRRR with the exception of the Asian model, which predicted marginally lower logRRR.

Proportion FRL. Proportion FRL was a significant predictor of logRRR in all models. In all but the Asian model, it was a relatively strong predictor second only to proportion White (Black, Latinx, NH/PI) or high school level (TomR). Proportion FRL was the strongest predictor of logRRR in the AI/AN model (sr_i = *-.191, or 19.1%)*. Higher proportions of students eligible of free/reduced lunch predicted lower logRRR in all models, except the Asian model. While the Asian model predicted a positive relationship between proportion FRL and logRRR, proportion FRL accounted for substantially less variance (sr_i = *.055 or 5.5%)* in logRRR relative to the other five models.

School level. Middle schools were significantly different from elementary schools in predicting logRRR in half the models (AI/AN, Asian, TomR). High schools were significantly different from elementary schools in all but one model (NH/PI). Other schools were significant in all models in which they were represented (NH/PI and TomR samples did not include any Other schools). In all 6 models, across all school levels, where school level significantly predicted logRRR, it predicted lower logRRR. This means that in all significant cases, the models predict that elementary schools will have larger gaps in suspension between students of color and White students as compared to middle, high and other school levels. As shown by the semi-partial correlations, the strength of school level as a predictor of logRRR is particularly notable at the high school level as compared to elementary schools. Indeed, high school level is the strongest predictor logRRR in the TomR model. In comparison to the entire set of predictors in the study, the consistency of school level as a predictor of school level racial/ethnic discipline gaps, both in terms of significance and directionality, was quite remarkable. Only the proportion of white students was a more consistent predictor of the dependent variable.

School locale. City schools were significantly different from suburban schools in predicting logRRR in all but one of the models (AI/AN). Town schools were significantly different from suburban schools in predicting logRRR in all models, and rural locales were significant in all but two models (Latinx, NH/PI). The semipartial correlations for city locales in all significant models are notably small. City locales (as compare to suburban locales) are also mixed in terms of directionality, with Asian, Black, and Latinx models predicting a positive relationship with logRRR and NH/PI and TomR predicting

the opposite. The models predicted that town locales would be associated lower logRRR in AI/AN, Latinx and NH/PI models as compared to suburban locales, while the other models predicted higher logRRR in town locales as compared to suburban locales. The amount of variance explained by the difference between town and suburban locales in the AI/AN (sr_i = -.102 or 10.2%*)* and NH/PI (sr_i = -.130 or 13%) models was notably large in comparison to other racial/ethnic groups and locales. Rural locales were a significant predictor of logRRR as compared to suburban locales in four (AI/AN, Asian, Black, TomR) of the six models. The amount of variance explained by the difference between rural and suburban locales in the AI/AN and Asian (sr_i = -.128 or 12.8% for both) models was notably large in comparison to other racial/ethnic groups and locales. Of the significant models, AI/AN, Asian, and Black predicted lower logRRR in rural vs. suburban schools, while the TomR model predicted marginally higher logRRR in rural vs. suburban schools.

Model Fit (R^2)

As indicated in Table 14, the Latinx model was by far the best fit (R^2 = .440), followed by the Asian model (R^2 = .326). Examining the regression results between the six models suggests that the higher R^2 for the Latinx and Asian models may correspond to a significant and substantial influence of proportion LEP. The AI/AN (R^2 = .265), Black (R^2 = .307), and NH/PI (R^2 = .297) models are similar in that proportion White and proportion FRL are the major contributors to R^2, various school locale and school level variables are relatively strong predictors of logRRR in each. When compared to the other models, TomR (R^2 = .176) model had a substantially lower R^2 value, which may be related to the observation that while proportion White and proportion FRL are strong

predictors of logRRR in the TomR model, their influence is much weaker in comparison to the other racial/ethnic groups. The R^2 calculations provide support for the primary hypothesis across all six models.

Summary of Notable Findings

The results of the quantitative analysis demonstrated the extent to which the proportion of White students was associated with the logged RRR, which compared the risk of suspension for students of color to the risk for White students within the same school. Control variables were also tested as part of the analysis, and included (a) student gender balance, (b) proportion of students with disabilities, (c) proportion of English learners, (d) proportion of students eligible for free or reduced lunch, (e) proportion of novice teachers, (f) number of school counselors, (h) school level, and (g) school locale.

The proportion of White students was the strongest overall predictor of logRRR across racial/ethnic groups. Furthermore, proportion White consistently predicted substantially higher racial/ethnic gaps in suspension, while controlling for the other variables in the model, all of which have been positioned in previous literature and theory as predictors of discipline outcomes.

The proportion of students eligible for free/reduced lunch was also a significant predictor of logRRR across all the models, although the effect of proportion FRL was relatively low in the TomR model and substantially lower in the Asian model. In all cases, proportion FRL predicted lower logRRR, suggesting that schools with higher poverty levels tend to have smaller racial/ethnic discipline gaps.

Results also provide strong evidence that as compared to middle, high and other level schools, elementary schools have larger suspension gaps between students of color

and Whites. This pattern was consistent for the significant coefficients across all racial/ethnic groups and all school levels, although the difference was most pronounced when comparing elementary to high schools. This suggests that there is something of a continuum from elementary – middle – high school, with the racial/ethnic suspension gap decreasing as one moves from primary through secondary school.

The proportion of LEP students was a substantial predictor of logRRR in only two models, Asian and Latinx. Given that Asians and Latinxs are the two largest LEP populations in U.S. schools (National Center for Educational Statistics, 2017), it might be expected that language status would have the most significant effect on school outcomes for those groups. Both the Asian and the Latinx models predict that higher proportions of LEP students will be associated with lower logRRR. This finding suggests that attending schools with less language diversity may place LEP populations at increased risk for suspension, or inversely, that attending linguistically diverse schools may serve as a protective factor against over-selection for suspension in LEP populations.

The results for school locale were more mixed overall, and more often than not, locale distinctions did not appear to have a substantial effect on logRRR. However, if we examine the cases in which locale was a substantial predictor, that is town schools in the AI/AN and NHPI models and rural locales in the AI/AN and Asian models, suburban schools appear to be associated with higher logRRR when compared to other non-city locales. Given the historical correlation between racial/ethnic diversity, poverty levels, and urbanicity, and it is unclear why no consistent, substantial differences between city and suburban schools were detected. However, it may suggest that the city and suburban locale designations used by the National Center for Educational Statistics are too “blunt”

in nature. That is, there may be a relatively high degree of variation *within* the city and suburban designations, with relative homogeneity within the town and rural designations. This would make it more difficult to detect any difference between city and suburban.

It is also notable that gender and disability status did not have a substantial effect in any of the models. Past research on the incidence of suspension has been virtually unanimous in finding males and students with disabilities tend to be suspended more often. The lack of effect in the current study may indicate limited variation between schools in gender balance and the proportion of disabled students. It clearly demonstrates that the predictors of the incidence of suspension (which students are suspended more often) are not predictors of the discipline gap between Whites and students of color.

Looking at the descriptive and inferential results together, we find that in the samples used in the current study, schools with fewer White students tend to have higher proportion FRL, higher proportion LEP and city locale designations. Results show that these schools also tend to have lower racial/ethnic discipline gaps for suspension, although the case for city schools is not particularly compelling. Nonetheless, overall these results suggest that schools exhibiting more poverty, more language diversity, and perhaps more urbanicity may contain protective factors associated with less racially/ethnically disparate suspension outcomes.

The next chapter positions these results in the context of existing literature and discuss what the results suggest for future research and school discipline policies and practices.

CHAPTER 5: DISCUSSION & IMPLICATIONS

A wealth of previous literature has documented differential disciplinary outcomes between students of color and their White counterparts across a variety of school contexts. Most previous research has understandably focused on Black students, as decades of research has found them to be disciplined more often and more harshly than any other racial/ethnic student group (Children's Defense Fund, 1975; Public Counsel Law Center, 2012). Studies have also found consistent over-selection of American Indians and Latinxs, though with less consistency (Brown & DiTillio, 2013; Peguero & Shekarkhar, 2011), while research has found that Asians tend to be under-selected for discipline in comparison to White students (Wallace et al., 2008; Rocque & Paternoster, 2011). Virtually no previous scholarship has examined discipline disparities in Native Hawaiian/Pacific Islander or students of two or more races. Racial/ethnic disparities have been linked to negative effects on the academic outcomes, social/emotional well-being and ultimately the life chances of the students of color that have most often been the target of disproportional discipline (Fabelo et al., 2011; Losen & Martinez, 2013; Wallace et. al., 2008).

As researchers have tried to explain racial/ethnic differences in school discipline, it has become clear that while student level variables, such as socioeconomic status and behavioral traits, are important predictors of student's risk of suspension (Raffaele Mendez et al., 2002; Rocque & Paternoster, 2011), these factors have been insufficient to

explain the presence of substantial racial/ethnic disparities (Fabelo et al., 2011; Wallace et al., 2008). Indeed, evidence has indicated that racial/ethnic disparities may be better explained by school level variables related to racial/ethnic composition, school contextual factors, and the disposition of school authorities (Saporu, 2012; Skiba et al., 2014). In particular, school racial/ethnic balance (i.e. % Black/White) has been shown to be one of the most powerful predictors of the incidence and severity of exclusionary discipline (Rocque & Paternoster, 2011; Skiba et al., 2014; Welch & Payne, 2010). However, it remains unclear how racial/ethnic balance affects the *magnitude* of racial/ethnic discipline gap between students of color and their White counterparts.

The typical approach to the study of racial/ethnic balance and has been to predict the discipline rates, or the odds of discipline for various racial/ethnic groups (usually Black, White, and Latinx) based on a set of predictors (including racial/ethnic balance) related to student risk using multilevel models to account for student, teachers, school level variables. In such studies, differential rates or odds of discipline for different racial/ethnic groups would be interpreted to indicate discipline disproportionality. As addressed in Chapter 2, this research has yielded a wealth of valuable information about what kinds of variables influence student level risk of discipline. The discussion in this chapter attempted to address whether the variables that best predict the student-level odds/risk/severity of discipline are the same variables that best predict school-level magnitude of racial/ethnic disparities.

Only a handful of studies have directly modeled metrics of racial/ethnic discipline disproportionality (Eitle & Eitle, 2004; Freeman & Steidl, 2016; Thornton & Trent, 1988). Results from the studies that have modeled racial/ethnic discipline disparities

directly (such as the relative risk ratios used in this study) have suggested that the factors that predict overall suspension rates or odds rates are not the same factors that predict racial/ethnic disparities (Freeman & Steidl, 2016; Thorton & Trent, 1989). However, this small body of work has substantial limitations; it examined only single state or district samples, analyses were limited to the Black-White racial/ethnic dichotomy, and there has been little examination of elementary schools. Careful reading of previous literature suggests that the racial/ethnic discipline gap may extend well beyond differences between Black and White students, and that the schools in which students of color are suspended most often do not necessarily exhibit the same characteristics as schools that tend to produce the largest racial/ethnic gaps. By extension, it is reasonable to assume that the kinds of interventions that may reduce the overall incidence or severity of discipline may be qualitatively different than those intended to reduce disparities in exclusionary discipline between Whites and students of color. Further, the limited research on school level racial/ethnic discipline gaps in suspension has not produced a clear answer to questions around what kinds of schools are most in need of interventions intended to reduce the very real harm associated with the disproportional application of suspension to students of color (Skiba et al., 2014).

This study was designed to add to the existing literature on racial/ethnic composition of place, school discipline, and racial/ethnic differences in exclusionary discipline. This study used bivariate regression with controls to examine a series of national samples to determine how racial/ethnic composition of place (proportion White) influenced the school-level discipline gap between six U.S. Census designated student

groups of color (American Indian/Alaska Native, Asian, Black, Latinx, Native Hawaiian/Pacific Islander, Two or more Races) and their White counterparts.

This chapter provides a review of the major descriptive and inferential findings with relation to previous research, conclusions based on the findings, and theoretical and practical implications regarding the issues raised in the research. This is followed by a discussion of the study's limitations and delimitations. The chapter concludes with suggestions for policy, practice and future research.

Discussion of Descriptive Findings

Racial/Ethnic Discipline Disparities

Descriptive results demonstrated that all student groups of color had mean RRRs levels above 1.00, indicating that across the large study samples, the incidence of suspension for all student groups of colors substantially exceeded that of White students within the same schools. This finding joins over 40 years of scholarship that has found that Blacks are disciplined more often and more harshly than White students, after controlling for the influence of a host of related variables (Children's Defense Fund, 1975; Fabelo et al., 2011; Public Counsel Law Center, 2012). It accords with past studies finding disproportional discipline of AI/AN students as well (Brown & DiTillio, 2013; Krezmien, et al., 2006). Substantial disparities in suspension were found between Latinxs and Whites as well, however the magnitude of over-selection was lower than was found for Blacks or AI/ANs. This finding is also in line with previous research on Latinx discipline disproportionality (Peguero & Shekarkhar, 2011; Skiba et al., 2011).

With regard to Asian students, the results of this study diverged from past research, which has consistently found Asians tend to be under-selected for discipline

compared to White students (Krezmien et al, 2006; Wallace et al., 2008; Rocque, 2010; Rocque & Paternoster, 2011). As such, it provided some of the first empirical support for the presence of racial/ethnic discipline differences between Asians and Whites. The present study's divergence from past literature on this point could be a consequence of the size and distribution of Asians populations in U.S. public schools. As shown in Table 10, only about half of the sampled schools over-selected Asians. This distribution of schools in which Asians are over-selected for discipline is much narrower than the other groups in this study (except Latinxs). Also, Asians attend suburban schools in much greater proportions than other racial/ethnic groups. Thus, samples in previous research may not have been large enough, distributed widely enough, or suburban enough to consistently detect disparities between Whites and Asians. For instance, Wallace and colleagues (2008) used a nationally representative sample, but analyzed only about 2500 Asian students in total. Krezmien and colleagues (2006) studied all the student in the state of Maryland. The over 7700 schools in the current study's Asian sample included over 870,000 Asian students in all 50 states and the District of Columbia.

This study also provided the first known large scale examination of the discipline outcome for NH/PI and TomR students in U.S. public schools. The magnitude of disparity for these two groups is noteworthy, as it is comparable to that of AI/AN and Black populations who have large, long-standing and well-documented disparities when compared to White students. Remarkably, TomR students had the highest mean RRR of any group in the analysis. As such, results of this study indicated that policy and reform related to the disparate disciplinary treatment needs to be expanded to account for disparities between Asian, NH/PI and TomR students and their White counterparts.

When combined with the findings that long-standing over-selection of AI/AN, Black, and Latinx students continues to persist, these new findings related to Asians, NH/PI, and TomR students suggest that inequitable discipline practices encompass even more students than previously documented. This, in turn, supported the conclusion that over-selection for suspension should be a point of concern in all school contexts where Whites and students of color attend school together. Put differently, the results of this study suggested that there is what might be referred to as a "baseline" level of over-selection for *all* students of color in most U.S. public schools. As will be discussed in more detail below, the presence of other factors (racial/ethnic balance, poverty, school level, language status) can serve to aggravate these baseline levels of disparity, producing even greater racial/ethnic discipline gaps.

While this study did not include data on racial/ethnic (mis)behavior patterns, past research has repeatedly concluded that patterns of misbehavior are similar between racial/ethnic groups (Finn et al., 2008; Huang, 2015). If one accepts that aggregate school level (mis)behavioral patterns are similar across racial/ethnic groups, findings like those from this study suggest that the over-selection of students of color for discipline is unjustified (Rocque, 2010, Rocque & Paternoster, 2011) because punishment would presumably be based (at least in part) on factors other than actual student behavior. That is not to say that widespread differential treatment is entirely unanticipated. Indeed, critical race theory has predicted such an outcome based on the position that racism is normal in American society and appears ordinary and natural to social actors (Delgado & Stefancic, 2000). The clear implication is that students of color would consistently face race-based assessment and interpretation of their behavior in U. S. public schools where

upwards of 80% of teachers and principals identify as White and middle class (U.S. Department of Education, 2012a, 2012b).

While this study found that students of color from all racial/ethnic groups face substantial over-selection for suspension across a wide variety of school contexts, the magnitude of disparity and the predictors accounting for disparity varied across groups. Results indicated that language status was a strong predictor of over-selection of Latinxs and Asians, whereas locale was much more important in predicting AI/AN, and school level best predicted TomR discipline disparities. This finding supported the suggestion within integrated threat theory (ITT) that racial/ethnic groups elicit different perceptions of threat, and thus different emotional and behavioral responses from Whites (Neuberg & Cottrell, 2002).

The Distribution of Racial/Ethnic Disparities in Suspension

Descriptive results added to past discipline literature through an examination of the distribution of schools that exhibit racial/ethnic disparities in suspension (See Tables 8 & 10). While past research has made it plain that African Americans are at increased risk of suspension in comparison to Whites across virtually all school contexts, large scale school-level comparisons across multiple racial/ethnic groups have not been forthcoming. As a result, it has been unclear whether only a small number of schools exhibited exorbitant racial/ethnic discipline gaps, or whether disparities were evident across a larger proportion of school contexts. This study provided evidence of the latter. Racial/ethnic discipline disparities appear to be pervasive and widespread for all students of color at the school level. AI/AN, Black, NH/PI, and TomR student groups were over-selected in the vast majority (>80%) of schools. Latinxs were over-selected in a strong

majority of schools (>60%), while Asians were suspended more often in over half (>50%). These findings confirmed that school level racial/ethnic discipline gaps are in line with those that have been documented on the student level for AI/AN, Black and Latinx students. Findings also provided new evidence that NH/PI and TomR students reflect similar distributional patterns of disparity across schools as AI/AN and Black students.

Looking at the magnitude and distribution of disparity across all six racial/ethnic groups revealed a two-tier pattern, where AI/AN, Black, NH/PI, and TomR students experienced school level disparities that are large in magnitude (3.00+ RRR) and widely distributed (>80% of schools), and Latinxs and Asians experienced substantially lower disparities across a smaller proportion of schools. Therefore, an implication that was drawn from this analysis is that race/ethnicity may influence discipline in different ways for Latinxs and Asians than it does for other students of color. While not tested in this study, the beliefs, attitudes, and dispositions of school authorities toward Latinxs and Asians may be different in ways that produce less racialized interactions or less robust perceptions of intergroup threat. For instance, levels of cultural threat perceived by predominantly-White school authorities from non-native English speaking students may not be as high as those based on misguided associations between Black male students and criminal behavior. Likewise, cultural stereotypes applied to Asians and Latinxs may elicit less racialized reactions than those applied at AI/AN, Black, NH/PI, and TomR students.

While the precise nature of the socio-cultural factors involved will require more focused research, the widespread nature of racial/ethnic discipline gaps nonetheless provided a clear exemplar of the level of "normalness" that inequitable application of

punishment has attained in the institution of U.S. public schools (Harris, 2012). Overall, the descriptive findings made it clear that racial/ethnic disparitieses do not reside within a few bad apples. It is worth reiterating that 9 out of 10 of the roughly 47,000 schools in their combined samples over-selected Black and TomR students for discipline, and that on average, for each White suspension in those schools, there were almost four Black and TomR suspensions. Thus, a crucial implication here is that, in the absence of reasonable justification, such widespread racial/ethnic disparities will likely demand more focused reform to prevent continued harm to students of color. This will be discussed below in more detail.

Discipline in Predominantly-White Schools

The descriptive results of this study added to previous scholarship on the intersection of racial/ethnic composition of place and school discipline through an examination of how predominantly-White status affected the racial/ethnic discipline gap. Very little research has been conducted on discipline in PWS (but see Rocque and Paternoster, 2011), and no known studies have directly addressed the magnitude of racial/ethnic disparity therein. As is evident from Table 9, this study found clear evidence that PWS exhibit larger gaps in suspension between students of color and their White counterparts attending the same schools. These findings suggested that attending majority White schools may expose students of color to more racialized structures of school discipline. The term "racialized" refers to differential discipline outcomes that could be reasonably attributed to the racial/ethnic composition of the school. Furthermore, greater risk suspensions for students of color in PWS is also likely to increase their exposure to

the negative consequences of exclusionary discipline (Fabelo et al., 2011; Losen & Martinez, 2013; Wallace et. al., 2008).

Previous scholarship has established a relationship between racial/ethnic balance and the odds/severity of student-level discipline (Skiba et al., 2014; Welch & Payne, 2010). This study provided complimentary evidence on the nature of the relationship between racial/ethnic balance and the magnitude of school-level racial/ethnic discipline disparities. Such evidence is likely to become increasingly relevant as students of color increase as a proportion of the total U.S. student population, and as they become increasingly dispersed into historically White school contexts (Colby & Ortman, 2014; Orfield & Frankenberg, 2014; U.S. Census Bureau, 2015).

Descriptive results related to schools' predominantly-White status were also relevant to theoretical arguments linking present and historical economic, social, legal, and educational arrangements to the enactment of systemic, institutional racism (Bonilla-Silva, 2010). As in some previous literature (i.e. Williams, 2015), the results of the present analysis suggested that the racial/ethnic milieu in U.S. public schools may reflect historical and present social arrangements that contribute to the enactment of racist discipline policies and practices. While not examined directly in this study, the widespread use of zero-tolerance discipline policies to suspend disproportional numbers of students of color, along with the lack of provision for the equitable application of discipline in federal education law, are just two glaring examples of what is meant by the above terms such as "legal arrangements" and "social arrangements." Although additional study is necessary, the clear implication was that the racism emanating from such arrangements may be part of what is reflected in racial/ethnic discipline differences

like those documented in the present study (Triplett, Allen, & Lewis 2014; Vavrus & Cole, 2002; Watts & Erevelles, 2004). Likewise, results of this study suggested that PWS are a potential site for the application of the kinds of dispositions/beliefs/attitudes that may seek the use of school discipline as a way to deny youth of color the full benefit of public education (Ladson-Billings & Tate, 1995; Welch & Payne, 2010).

Discussion of Inferential Findings

Proportion of White Students

The regression models for all six racial/ethnic groups demonstrated that the proportion of White students was the most powerful predictor of school-level discipline disparities in the samples analyzed. This reinforced the salience of race/ethnicity in discipline (in)equity, and suggested that race/ethnicity operates at the school level across all racial/ethnic groups within a broad range of school contexts. This finding suggested that all other things being equal, students of color will be more overrepresented in suspension in schools with more White students. By extension, in schools with more White students, students of color may disproportionately experience the negative academic and social-emotional consequences of suspension.

A substantial body of extant literature has examined how racial/ethnic balance affects the racial/ethnic odds and severity of school discipline, however previous literature has used the proportion of Black students to operationalize racial/ethnic balance. This literature has shown that larger populations of Black students are related to increased student-level risk and severity of discipline for students of color (Saporu, 2012; Skiba et al., 2014; Welch & Payne, 2010). Although not directly analogous to the present study, which used the proportion of *White* students to predict the magnitude of

racial/ethnic disparity at the school level, this study's results did support previous findings to the extent that increased student-level odds of suspension are connected to increased school-level racial/ethnic gaps in suspension.

A handful of studies have directly modelled the magnitude of racial/ethnic disparity between students of color and White students using some variation of RRR, just as in the present study. Thornton and Trent (1988) found a negative correlation between the proportion of Black students and Black vs. White disproportionality, as would be predicted by the present study. However, Eitle and Eitle (2004) found that the size of the Black population was not related to higher disparities between Black and Latinx vs. White students. The findings of the present study directly contradicted this finding. More recently, Freeman and Steidl (2016) operationalize racial/ethnic balance using a Black-White *exposure index*, which measured school diversity and the potential or interracial interaction within school districts. The authors were primarily interested in how de/resegregation affected discipline outcomes. They found that more segregated schools were associated with lower racial/ethnic discipline disparities. Their models predicted that increasing school Whiteness would produce lower racial/ethnic discipline gaps, a prediction that was contradicted by the present analysis.

While the findings of Freeman and Steidl (2016) may seem to be directly at odds with the evidence in the present study, it is entirely plausible that the proportion of White would predict larger racial/ethnic discipline disparities (as was found in the current study), and that increasing exposure between White and students of color (especially in the context of de/resegregation) could also contribute to over-selection of students of color for discipline within the same schools. Indeed, the theoretical grounding of the

present study would predict such a dynamic. For instance, critical race theory may point to the aforementioned mix of social, residential, and educational arrangements present in historically White school contexts to help explain why greater proportions of White students would predict higher racial/ethnic discipline gaps. At the same time, intergroup threat theory would likely point out that as populations of students of color were segregated out of predominantly-White schools, school authorities would perceive less of a threat to school culture and norms and may relax social control measures such as discipline against students of color. I return here to the notion of a baseline school-level disparity, as predicted by the proportion of White students used the present study, and potential aggravating factors such intergroup threat based as increasing populations of color, as predicted by the proportion or students of color (Welch & Payne, 2010).

Control Variables

Proportion of students eligible for free/reduced lunch. Regression analysis revealed a negative association between the proportion of students eligible for free/reduced lunch and the magnitude of racial/ethnic discipline disparities in all models, except that of Asians. The Asian model revealed a statistically significant, but substantively small positive association. This finding indicated that outside of Asians, students of color are given fewer suspensions in comparison to their White counterparts in schools with higher mean levels of student poverty. This suggested that schools positioned in areas that have higher income levels may tend to over-select students of color for discipline. While this finding may not be a complete surprise given the historical association between race/ethnicity and poverty status in the U.S (DeNavas-Walt &. Proctor, 2015), it does appear to qualify previous literature on the relationship

between socioeconomic status (SES) and discipline, which has largely concluded that lower SES at the student level predicted a greater likelihood of suspension.

Numerous studies on the odds or incidence of student-level suspension have found that lower income students are suspended more often than wealthier students (Nichols, 2004; Raffaele Mendez et al., 2002; Sullivan, Klingbeil, & Van Norman, 2013). A much smaller body of literature has examined how student income and poverty levels affect racial/ethnic discipline disparities. Thorton and Trent (1988) used median family income of the census blocks surrounding sampled schools as a proxy for school income levels. They found that higher school income levels correlated with higher Black-White suspension disproportionality. More recently, Freeman and Steidl (2016) used FRL status in a multilevel model to predict that schools with higher proportions of students from low income families have smaller discipline gaps between Black and White students. The findings from the present study are in line with this smaller body of work.

Thus, while lower SES has been associated with increased individual-level odds of suspension in past research (Nichols, 2004; Raffaele Mendez et al., 2002; Sullivan et al., 2013), lower SES was associated with higher racial/ethnic discipline gaps in this study. This finding suggested that while poverty may function as an aggravating circumstance when examining the likelihood or severity of discipline for individual students, such that students of poverty would tend to be overselected for suspension compared to their wealthier counterparts, attending a school with higher poverty levels may also serve as a protective factor when examining the kinds of schools that produce the widest suspension gaps between Whites and student groups of color. This finding provided additional support for the emerging recognition that the variables that predict

student-level odds of discipline are different than those that best predict racial/ethnic disparities in discipline at the school level (Freeman & Steidl, 2016).

School level. The results of this study indicated that the elementary school level is associated with larger racial/ethnic gaps in suspension as compared to middle, high and other school levels. This pattern was remarkably consistent and held true across all six models except for NH/PIs, although both middle and high school were not both significantly different from elementary schools in all models. The strength of school level as a predictor of logRRR was particularly notable at the high school level as compared to elementary schools. Indeed, high school level was a stronger predictor logRRR than the proportion of White or FRL eligible students in the TomR model.

None of the previous research that has directly examined school level disparities (RRR) included elementary schools in the analysis (Eitle & Eitle, 2004; Freeman & Steidl, 2016; Thornton & Trent, 1988), but studies have looked at the relationship between school level and the odds of student-level suspension of office referrals. Skiba and colleagues (2011) found that Blacks were far more likely to receive office referrals in middle schools as compared to elementary schools, while Latinxs were overrepresented in office referrals at the middle school level, but underrepresented at the elementary level. Butler (2011) found that elementary level students had significantly higher risk of suspension and expulsion than secondary (middle and high school) school students. The present study joined Butler (2011) in indicating that elementary schools appear to be an important site for the study of school discipline. This challenges the notion that middle and high schools are the areas of greatest concern (Eitle & Eitle, 2004; Losen & Gillespie, 2012). The results of the present study also suggested that elementary schools

are an important site not only for increased risk of suspension for students of color, but for the application of racially/ethnically disproportional patterns of exclusionary discipline. Therefore, a paradigm shift in the conceptualization of the ways that school discipline plays out across school levels may be necessary to explain why elementary schools would tend to over-select students of color for suspension so consistently.

Butler (2011) suggested that elementary schools may use suspension and expulsion more often because they tend to have limited options for temporary removal (such as in-school-suspension) of disruptive students. Given the present study's finding that race/ethnicity were implicated in differential suspension patterns across school levels, another possibility for the interpretation of this finding emerged from the current analysis. Past literature has found that racial/ethnic disparities are higher in subjective offense categories such as "disruption" or "disrespect" (Losen & Martinez, 2013; Skiba et al., 2002). To the extent that serious, objective infractions (i.e. weapons, fighting, drugs & alcohol) are less likely to occur at the primary school level, elementary schools may exhibit larger racial/ethnic discipline gaps because a higher proportion of disciplinary incidents at the elementary level may involve subjective offenses. The implication here again is that elementary-level school authorities may interpret the behavior of students of color in relation to negative stereotypes and biases, or they may feel compelled to discipline students of color as a means of preserving what they perceive to be the norms of their schools or classrooms. More disciplinary incidents based on subjective offenses may offer an avenue through which to apply negative attitudes and perceptions of threat more often in elementary school contexts.

Gender balance. Numerous studies have found that gender is a strong predictor of discipline outcomes (Fableo et al., 2011; Finn & Servoss, 2013; Rocque & Paternoster, 2011). Most, though not all (Butler, 2011), have found that males are suspended more often than females. No known research has examined the relationship between gender balance and school level racial/ethnic discipline gaps. The present study found that while school-level differences in gender are a significant predictor logRRR in some racial/ethnic groups within large samples, gender differences do not generally account for a substantial amount of the variation of school-level disparities in suspension. This result was surprising given the demonstrable salience of gender in student-level analysis of discipline. It also provided an additional example to support the emerging conclusion that different factors best predict of student level odds of discipline and school-level discipline gaps.

Disability status. Disability status has been shown to predict suspension rates, with student disability being positively related to suspension rates (Krezmien et al., 2006; Skiba et al., 2002; Wallace et al., 2008). In addition, like the present study, Freeman and Steidl (2016) modeled the relationship between the proportion of students with disabilities and the Black-White discipline gap. They found that disability was statistically significant, but a weak predictor of substantive differences in Black-White discipline. The present study arrived at the same conclusion for the Black-White discipline gap, and provided evidence of a similar result for the other racial/ethnic groups in the analysis. Thus, as was the case for gender, this study found that although disability is a substantial predictor of the incidence/risk of suspension, it does not appear be a strong predictor of racial/ethnic discipline gaps in suspension.

Limited English proficiency. As discussed in Chapter 3, there has been very little research pertaining to language status and school discipline. A previous study using 2011-2012 Office of Civil Rights data found that LEP students were slightly under-represented in the incidence of exclusionary discipline in Arkansas, but that Latinx LEP students were overrepresented when compared to Whites (Anderson & Ritter, 2017). Monroe (2013) studied discipline patterns in a suburban school in the US south, and found that discipline disparities for Asian and Latinx students were related to language status. The present study found that LEP is a strong predictor of lower logRRR for both Asians and Latinxs. LEP was not a substantial predictor of logRRR in the other racial/ethnic group models. These results were directly in line with the findings of Monroe (2013). This result provided yet more evidence that different factors may be important in the prediction of different racial/ethnic groups. It was noteworthy that Asians and Latinxs were the two largest ELL populations in U.S. schools (National Center for Educational Statistics, 2017). Given the role of language as a cultural marker and public debate related to immigration and English-only laws, this finding suggested that language may be another axis of threat upon which dominant social groups apply social control measures to non-native English speakers in an attempt to preserve language and cultural dominance.

Novice teachers. Studies have found that more experienced teachers tend to suspend more students of color (Eitle & Eitle, 2004; Raffaele Mendez et al., 2002), and staff experience levels vary across educational contexts (U.S. Department of Education, 2009b). An implication of this scholarship was that veteran teachers may take a more "no-nonsense" approach to school discipline, and that veterans may feel more compelled

to preserve their conceptions of school norms and culture through the use of school discipline as a means of social control as applied to students of color. In more recent scholarship, Freeman and Steidl (2016) modeled the relationship between the proportion of first year teachers and the Black-White and Latinx-White discipline gap. They found that the proportion of 1st year teachers was statistically significant, but a weak predictor of substantive differences in racial/ethnic discipline gaps. The present study found the proportion of novice teachers to be amongst the weakest predictors of logRRR, which supports the conclusion of Freeman and Steidl (2016). It is plausible that the proportion of 1st and 2nd year teachers is a poor proxy for teacher experience levels, or that influence detected in previous studies was associated with the proportion of veteran rather than novice teachers.

Full time counselors. Past research has found that referrals to the school counselor for behavioral concerns and disruptive behavior may contribute to differential discipline outcomes for Black students (Adams et al., 2007; Bryan et al., 2012). However, the present study found full time counselors to be a marginally substantive predictor of logRRR in all models except AI/AN. As explained in Chapter 3, this abnormally large coefficient for counselors is assumed to be a consequence of an interaction effect between full time counselors and town/rural school locales, in which AI/AN are overrepresented.

School locale. Research has shown that city schools serve higher percentages of students of color (U.S. Department of Education, 2010), and that urbanicity influences school discipline, primarily because large, urban districts tend to have higher incidence of suspension (Skiba et al., 2002; Wallace et al., 2008; Welch & Payne, 2010). Saporu

(2012) found that urbanicity does not predict risk of discipline for Black and Latinx students. However, no known studies have examined how locale influences the magnitude of racial/ethnic discipline disparities. The present study showed notably mixed and weak substantive results for school locale. Given the demonstrable influence of the proportion of White students and FRL, the limited predictive power of locale was surprising. One might have expected suburban schools to be strongly correlated with higher racial/ethnic disparities since they have more White students and fewer low-income students, both of which were found to be associated with larger racial/ethnic discipline gaps. Overall, results suggested that the suburban locale designation may predict slightly higher logRRR, however the strength of its influence is notably underwhelming. It is possible that the locale designations used by the National Center for Educational Statistics reflect the increasing overlap between traditionally defined city/suburban/rural contexts (Milner, 2012). Locale results might also have reflected growing numbers of students of color in rural and suburban school contexts (Colby & Ortman, 2014; Orfield & Frankenberg, 2014; U.S. Census Bureau, 2015).

Delimitations

This study was delimited by the following characteristics:

1. The data are limited to school level metrics and do not include observations of individual students, school authorities, or disciplinary incidents. Therefore, results may have limited application to individual students, teachers, school authorities, or disciplinary incidents.
2. School discipline data represent total incidence of school discipline. They do not differentiate between multiple incidents of discipline applied to a single

student (duplicated counts) and discipline applied to multiple students (unduplicated counts). Therefore, results may have limited generalizability to school contexts with exreme numbers of duplicated counts of suspension.

3. The data do not include metrics related the racial, ethnic and cultural dispositions of students, school authorities and other school-based actors. Therefore, implications and conclusions related to racial, ethnic and cultural dispositions of students, school authorities and other school-based actors should be viewed with caution.
4. It was not feasible to cross check data with the written transcripts that detail each disciplinary incident. Therefore, the metrics used to derive relative risk ratios may have been different to the extent that cross checking of individual disciplinary incidents may have produced different incident counts.
5. In using White students as a comparison group, it was impossible to determine if RRRs and logRRRs above 1.0 indicated an over-selection of students of color, or an under-selection of White (see Hosp & Reschly, 2003, p. 70). Based on past literature it is most likely over-selection of students of color.
6. This study used school-level data as reported by schools to the Office of Civil Rights (OCR). Schools across the nation presumably define, code and report incidence of school discipline (i.e. suspension, expulsion, referrals to law enforcement) in different ways. This variation would have influenced the degree of correlation between study variables to the extent that such variation in defining, coding and reporting prevented the statistical model(s) from grouping similar school discipline incidents in a consistent manner.

7. The data did not include metrics related to student (mis)behavior rates. It is theoretically plausible that observed differences between the proportion of discipline applied to a racial/ethnic group and that same racial/ethnic group's proportion of the school population could be caused by variations in behavior between racial/ethnic subgroups. However, research specifically examining the predictors of student (mis)behavior has found virtually no relationship between level of misbehavior and student race or ethnicity (Finn et al., 2008). Even studies that have documented variance in behavior between student racial/ethnic groups have concluded that differences in behavior are insufficient to explain overall racial/ethnic differences in school discipline outcomes (Kupchick & Ward, 2011; Rocque, 2010; Wallace et al., 2008).

Limitations

Schools were excluded from the analysis where the number of White students, the number of students classified within racial/ethnic subgroups, or the incidence of suspension made it impossible to provide stable estimates of relative risk. As such, schools with the most extreme disparities between Whites and students of color on both the high and the low end of the distribution were likely excluded. Therefore, results may not be generalizable to schools with extremely small numbers (< 10) of racial/ethnic groups, or schools with no reported instances of suspension during the 2013-2014 school year.

Recommendations for Policy & Reform

This study provided much needed evidence regarding school-level discipline patterns. It did so using large national samples, and provided comparisons between all

U.S. Census designated racial/ethnic groups and their White counterparts within the same schools. Previous studies of student-level differences in school discipline have provided a wealth of empirical and interpretive information, much of which has informed this study. However, it can be argued that school-level predictors become increasingly relevant within the context of potential intervention and reform, as it is likely untenable to plan scalable reforms on an individual, case-by-case basis. This study proceeded with the recognition that school-level predictors might help reveal the most critical areas to focus future study and reforms intended to reduce the harm of racial/ethnic discipline disproportionality. Educational actors have become increasingly aware of discipline issues, and policymakers have increasingly begun to take action intended to produce more equitable outcomes. While recent inquiries into racial/ethnic discipline disparities by the U.S. Departments of Education and Justice (U.S. Department of Education, 2016a) were both sorely needed and long in coming, intervention in individual, egregious cases based on citizen-generated civil rights complaints may not be sufficient to produce more equitable discipline outcomes on a scale that is commensurate with the scale of the problem as documented in the current analysis. Furthermore, there has been no indication that the ability of Obama-era officials to recognize and act upon (in however limited a fashion) inequitable school discipline practices will be continued in present or future administrations (Green, 2017). Therefore, the ability to predict what kinds of schools exhibit larger racial/ethnic discipline gaps has provided important contextual information to reformers and policymakers as they decide how to spend scarce education resources.

This section outlines three potential reforms based on the results of the analysis. These reform suggestions attempted to address the tiered character of discipline

differences in a way that addressed both structural, macro level factors and individual/micro level predictors of disproportionality. As always, decisions regarding school discipline policy must involve multiple stakeholders (particularly students of color, their parent(s)/guardian(s), their community leaders), and be based on the specific characteristics of the school community. As such, reforms below represent a broad sketch based on the results of this study and current literature and theory.

Federal Discipline Accountability

Given their magnitude and extent,racial/ethnic discipline disparities must be considered a civil and human rights issue. Students of color must not be asked to wait under constant duress as researchers and policymakers parse the details of how and why students of color are constantly over-exposed to suspension, and the attendant risks of academic and social-emotional damage, at rates much higher than their fellow White students in the same schools. In order to minimize future harm, a system of discipline accountability is proposed, built around existing federal academic accountability structures, which ostensibly conditions federal education dollars on meeting academic yearly goals (Triplett et al., 2016). In many ways discipline accountability is what might be termed “triage intervention” in that it represents a way to prevent ongoing, irreparable harm to students of color in American schools. This proposal is not a suggestion to halt efforts related to consciousness raising, the attenuation of racial/ethnic biases and stereotypes, or the implementation of more equitable discipline policies and practices like SWPBIS or restorative practices. While critical, such efforts are likely to involve long-term, intensive development and reform efforts. Discipline accountability is intended to prevent harm now as other kinds of intervention proceed.

Existing federal academic accountability is maintained largely through the Every Student Succeeds Act (ESSA) (2015) formerly known as No Child Left Behind, the explicit purpose of which is to "provide all children significant opportunity to receive a fair, equitable, and high-quality education, and to close educational achievement gaps" (ESSA, 2015, 20 U.S.C. § 1001). While previous versions of the ESSA have focused almost exclusively on academic standards, the most authorization has taken important steps toward school discipline reforms by acknowledging the presence and potential harm of excessive use of exclusionary discipline like suspension. It also requires the previously optional collection discipline data as part of state annual report cards. Thus, the ESSA provides an acknowledgement of the relationship between discipline and achievement, which school discipline scholars have referred to as "two sides of the same coin" (Gregory et al., 2010, p. 1). These steps have introduced the potential for discipline outcomes to be integrated into federal accountability structures previously reserved exclusively for results related to academic achievement and attainment. However, despite the progress made with the ESSA, it remains unclear how state and local discipline outcomes will be integrated into federal accountability structures, there is no conditioning of federal dollars on discipline outcomes, and there is no provision made for support or intervention to help schools move toward more equitable and less harmful discipline practices (Triplett et al., 2016).

Amending ESSA to provide for a system of discipline accountability may be the most efficient and comprehensive way to implement some measure of oversight, as a means of preserving the civil and human rights of students of color in U.S. public schools. Previous scholarship has provided detailed suggestions for reforming ESSA

(Triplett et al., 2016; U. S. Department of Education, 2010), specifically the inclusion of specific allowances for federal intervention in school discipline based on *disparate impact* legal analysis (Triplett et al., 2016). Disparate impact refers to a situation where seemingly neutral policies and practices produce disproportionate and unjustified effects. It can be contrasted with disparate treatment, which refers to instances of intentional discrimination. Disparate treatment has historically been difficult to prove in school discipline court cases (Advancement Project, 2005). Under the Obama administration, the U.S. Departments of Education and Justice pursued intervention in discipline cases based on a disparate impact legal frameworks (U.S. Department of Justice & Department of Education, 2014), however such intervention is unlikely under the Trump administration (Green, 2017). The inclusion of language explicitly allowing intervention based on disparate impact into the ESSA would preserve the ability to intervene in cases of racialized and unjustified school discipline, even if future leaders do not consider civil rights enforcement important, or until federal, state, or local legislative bodies produce legislation specifically disallowing it.

The current proposal of discipline accountability in schools aligns with previous work in social psychology and legal studies, which has suggested that the unwanted and often unconscious application of negative attitudes and biases toward marginalized groups can be attenuated when decision makers (such as school authorities) are held accountable for their actions (Casey, Warren, Cheesman, & Elek, 2012; Kang et al., 2012; Lerner & Tetlock, 1999; Reskin, 2005). In one of the few studies on the subject in the field of education, Pit-ten Cate, Krolak-Schwerdt, and Glock (2015) examined how accountability measures affected teachers' implicit racial bias, decision making, and

judgments of student achievement. Teachers read a series student profiles that varied by race and responded to questions about the student's future achievement. Results demonstrated that perceived accountability increased teachers' accuracy in predicting students' future achievement through a reduction in racial/ethnic biases. However, in a 6-month follow-up, teachers' biases reappeared although not to the same extent as initial measurements. While the Pit-ten Cate et al (2015) study examined perceptions of student academic abilities, similar results are plausible for school authorities' perceptions of student (mis)behavior. This emerging line of research highlights the potential of accountability measures as a component of intervention in disparate suspension outcomes and emphasizes the need for accountability provisions that are consistent and intensive in nature.

For educational actors on the state, district, and school level, a system of discipline accountability will likely require additional tracking and analysis of school discipline outcomes. As is the case with federal accountability structures related to academic achievement, educational entities will need to measure and track how schools are doing for each group of students on each of the indicators in the nascent discipline accountability system. They will also need to identify any school that is consistently overselecting any group of students for discipline and produce plans for support and improvement. The ESSA leaves many key accountability decisions up to states, such as what to measure, how to communicate how schools are doing on those measures, how to identify when a school needs to take action to improve for any group of students (Smith & Lowery, 2017). Thus, "the decisions states make as they put together their ESSA plans

signal whether they will continue on a harmful path of masking inequities or whether they will tackle them head on (p. 7).

Separating Subjective and Objective Disciplinary Processes

The second reform suggestion is to create some separation between the processes involved in handling subjective and objective offenses. While there is no clear line between the two, subjective offenses are relatively ordinary student behaviors not traditionally understood as criminal, such as insubordination, disrespect, or disruption that are generally unrelated to school safety (Advancement Project, 2000; Gregory et al., 2010). Objective offenses are related to criminal behavior or school climate/safety, such as weapons, drugs, alcohol, tobacco, threats, and fighting. Scholars have long examined subjective offenses as a potential explanation of racial/ethnic discipline differences. Research has found that students of color are often sent to the office based on subjective offenses (Gregory et. al., 2010), and that racial/ethnic discipline disparities are more likely to be found in subjective offense categories (Losen & Martinez, 2013). Skiba and colleagues (2002) found that students of color were referred to the office more often for subjective offenses, while White students were referred more often for objective violations. Fabelo and colleagues (2011) revealed the Black students were often significantly less likely than White and Latinx students commit serious (typically objective) offenses trigger mandatory disciplinary consequences.

Part of the issue with subjective offenses is that they are "subject to the discretion of school employees" (Fabelo et al., 2011, p. 46). There is mounting empirical evidence that the questionable attitudes and dispositions of school authorities factor into the production of racial/ethnic variation in discipline outcomes. Research has suggested that

Black students are frequently stereotyped as “troublemakers” (Fenning & Rose, 2007; Okonofua & Eberhardt, 2015). Indeed, school authorities appear to perceive the behavior patterns and academic abilities of students of color more negatively based on cultural characteristics like dress, speech patterns, and patterns of social interaction (Neal et al., 2003). Therefore, school discipline represents an arena in which pervasive discriminatory ideologies about students of color can “leak” into discipline outcomes, contributing to disproportional punishment and exclusion (Butler, Robinson & Walton, 2014; Gregory & Weinstein, 2008; Watts & Erevelles, 2004).

The intention of this proposal to separate objective/subjective offenses is to eliminate some of the discretion involved, and to provide more oversight in cases of discipline based on subjective offenses. It is recommended that subjective offenses be flagged, documented, and processed in a different manner. This would likely require new ways of documenting and tracking disciplinary incidents, such as redesigning office referral forms and implementing computer tracking systems that differentiate between objective and subjective offenses. As such, school authorities at the state, district, and school level would likely need to clearly delinaeate what is meant by nebulous subjective offenses such as disrespect and insubordination, while also clearly delineatingt what might be considered subjective vs. objective offenses. It is imperative that these efforts be inclusive of all educational stakeholders (teachers, students, parents, school professionals, community members). An inclusive process presents the best opportunity of producing genuine dialogue around the culturally conditioned interpretation of student behavior, and of producing unbiased parameters around subjective offenses that can limit the application of racial/ethnic bias and stereotypes.

New, differentiated subjective offense processes could be modelled after required meetings for the long-term exclusion of students with disabilities. For instance, when students identified as disabled face long-term suspension, a team of staff members conduct what is often referred to as a "manifestation" meeting, in which they determine the extent to which a child's disability contributed to the violation (Maryland Disability Law Center, 2015). In terms of subjective disciplinary offenses, the intention of a similar meeting would be to involve additional stakeholders in the process to provide a second look at the violation outside of the pressure of the actual incident. Reasonable choices for those additional stakeholders include school counselors (Kyle & DeVoss, 2015), parent oversight boards (Zinth, 2005), or school-based civil rights coordinators (CRCs). Civil rights coordinators are employees designated to coordinate schools' compliance with federal civil rights law. Enacting more rigorous review of subjective offense suspensions might reduce the number of subjective disciplinary incidents elevated to the office. If future research indicates that a higher proportion of elementary school disciplinary incidents are based on subjective offenses, this reform might be particularly helpful in addressing the surprisingly large and consistent discipline gaps at the elementary level revealed in this study.

Slow Discipline & Restorative Practices

The unconscious nature of implicit bias along with the high-stress, moment-to-moment nature of classroom interactions suggest that the role of personal bias in disciplinary decisions could be diminished by what might be termed "slow discipline." Slow discipline emphasizes taking time to fully understand aggravating classroom interactions. Slow discipline also prizes attention to the presence of potential biases and

purposefulness in avoiding the application of personal bias in the interpretation of another's behavior. This kind of attention further implies the need to alter common notions about school discipline—away from dispositions that see school authorities as the sole decider of disciplinary outcomes and toward greater emphasis on reflection and self-awareness for school authorities.

Reflective and self-aware school authorities have the potential interrupt the potentially influential role of negative past experiences with students of color by recognizing that each student, each day and each incident need not be influenced by past experiences. As Elsa Barkley Brown (as cited in Lemert, 2010) has stated, "all people can learn to center in another experience, validate it, and judge it by its own standards without need of comparison or need to adopt that framework as their own…one has no need to 'decenter' anyone in order to center someone else" (p. 551). In terms of school discipline, the implication is that the process of racialized school discipline could be interrupted if predominantly-White school authorities were able to "decenter" their own cultural identities as a means of understanding the standpoint of their students of color. The related practice of critical reflection (Howard, 2003) also seeks to address issues of equity in the classroom and might serve as a model for staff development efforts specifically aimed at implicit bias.

Restorative justice discipline models (Costello et al., 2010; Schiff, 2013) also emphasize a slower discipline process. In restorative justice models, disciplinary events often trigger the use of a dialogue circle that typically includes teachers, students and a third-party facilitator. Circles allow responsibility for disciplinary incidents to be shared between students and school authorities and can highlight how disciplinary incidents are

escalated by external factors, like stress, individual circumstances or the presence of negative stereotypes or deficit ideologies. This kind of approach values student voice, removes the role of teachers as the sole arbiter of misbehavior, and promotes dialogue about the root causes of problem behavior. Research has found that restorative practices can decrease the use of punitive school discipline (Lewis, 2009), more positive relationships between teachers and their diverse students, and fewer discipline referrals for Latinx and African American students (Gregory et al., 2016). Wider implementation of restorative practices may also help clarify how more restorative discipline consequences (with fewer negative effects on students) are portioned out to various student racial/ethnic groups, particularly in schools with higher proportions of White students, where research (Payne & Welch, 2010Welch & Payne, 2010) has suggested the presence of a rationing effect, whereby punitive consequences are meted out to students of color, and restorative practices (i.e. referrals to medical professionals) are reserved for White students.

Future Research

Generally, this study highlighted the need for more research on the disciplinary practices in elementary schools, which may help to qualify prevailing notions about school level and school discipline that position secondary schools as the areas of primary concern (Eitle & Eitle, 2004; Losen & Gillespie, 2012). Such notions may have inhibited the kind of close examinations needed to explain exactly why elementary schools tend to magnify the risk of suspension for students of color as compared to Whites. As noted above, analysis of patterns of subjective offense incidents is likely a good starting point.

Qualitative studies of disciplinary norms, particularly observation and interpretation of elementary disciplinary incidents would be valuable as well.

Additional research is also needed on TomR and NH/PI student populations, as the new evidence presented here positions these as critically under-researched student groups in the discipline literature. TomR students exhibited alarmingly high rates of suspension compared to White students. These students are numerous (over 17,000 schools had 10+ TomR students) and widespread (their sample included more schools than AI/AN, Asian and NH/PI combined). Multiracial students also exhibit wide in-group variation (Pew Research Center, 2015). Without additional in-depth study that accounts for in-group variation, it will be difficult to generalize about TomR students, and thus difficult to determine effective intervention to reduce the startlingly high risk of over-selection for suspension found in this study.

Results also call for comparison studies between Latinx, Asian and the remainder of the racial/ethnic groups in order to analyze why Asian and Latinx populations have much smaller racial/ethnic discipline gaps and are over-selected in substantially fewer schools.

Future studies need to provide additional parsing of racial/ethnic group data collection and analysis. For instance, South Asians are quite distinct from East Asians in terms of language and migration status, and socio-demographic patterns (Hsin & Xie, 2014). As mentioned above, within the group identifying as TomR, racial/ethnic background varies widely (Pew Research Center, 2015). Given that these kinds of characteristics are closely related to the substantive predictors of Asian-White discipline

differences, further division of racial/ethnic groups in data and research would likely increase understanding of disproportionate educational outcomes.

Future studies of racial/ethnic discipline differences should incorporate metrics of racism, bias and racial threat into analyses (see Okonofua & Eberhardt, 2015; Trachok, 2015). Past studies on the disposition of school authorities have focused on punitiveness (Skiba et al., 2015), but the theoretical orientation of this study suggested that overly-punitive attitudes toward students of color have their origins in racism and racial/cultural threat. However, just the presence of disparities may suggest, but does not explicitly implicate racism or bias in the production of disparities (Williams, 2015). Likewise, the simultaneous presence of more black students, and higher Black-White discipline gaps suggests, but does not implicate racial threat (Welch & Payne). Without empirical evidence that school authorities apply bias and negative attitudes to disciplinary decisions, the ability to intervene on behalf of students of color may be constrained.

Future research should consider the inclusion of the proportion of White students into a more comprehensive metric of school Whiteness (Williams, 2015) that might include the racial/ethnic traits of teachers and school leadership at least, and possibly some examination of the racial/ethnic balance of the communities from which student bodies are drawn. A more comprehensive metric of school Whiteness may reveal how a culture of Whiteness influences the decisions of school-based actors in educational contexts currently or historically associated with White culture. It may also clarify how racial/ethnic composition of place affects discipline, along with other important educational outcomes related to attainment.

Summary and Conclusions

Looking at the full set of predictors in the present study with reference to previous literature led to the conclusion that predictors of student level racial/ethnic differences in the odds, likelihood, or severity of discipline are not necessarily strong predictors of school-level racial/ethnic disparities. Given that only a single study (Freeman & Steidl, 2016) has suggested this previously, the present study provided critical supporting evidence. This is a meaningful distinction because it is relevant to what kinds of interventions are likely to be effective and where to focus intervention and reform. Intervention intended to address student-level disparities may need to be different than those focused on school level disparities. This suggests that reform may need to multifaceted, and tiered in such a way that enables it to address both student and school level disparities.

A tiered response to discipline issues may also help in differentiating the policies and procedures for addressing objective and subjective offenses to prevent the introduction of bias and stereotypes into discipline outcomes. Emerging research has suggested multi-tiered frameworks, primarily in the context of school-wide positive behavioral support and intervention strategies, with more intensive supports offered when more general strategies fail to resolve behavioral problems (Gregory et al., 2017). These strategies, along with more intensive use of data and more restorative discipline strategies (Gregory et al., 2014) hold promise for addressing racial/ethnic suspension gaps. However, tiered strategies that fully appreciate the nuances between subjective vs. objective offenses and student vs. school level disparities have yet to emerge. As such, the proposal of a system of discipline accountability forwarded in this study represented

the first steps toward intervention intended to honor the legal, civil and human rights of students of color by minimizing ongoing harm now, even as more equitable policies, practices, and dispositions are pursued with all deliberate haste.

This study has also made it clear that different variables predict disparities across racial/ethnic groups. As mentioned above, LEP appeared to be empirically and theoretically relevant to Latinx and Asian populations. FRL eligibility was a strong predictor of discipline gaps across all groups, but was found to have little effect on Asian discipline. Similarly, locale distinctions were highly predictive for AI/AN, while only marginally useful in other group models. While this study began to address calls for discipline research to move beyond the Black-White and sometimes Latinx and AI/AN paradigm (Freeman & Steidl, 2016), it also implied that analyses based on only a limited number of racial/ethnic subgroups may fail to provide the nuance necessary to inform focused intervention. Given that the present study is a school level analysis, suggest that the variations observed between groups reflected some degree of true difference between schools, and that those differences may have influenced how race/ethnicity intersects with discipline policy and practice.

To summarize, the results of this study provided evidence that school-level racial/ethnic disparities are widespread and pervasive in U.S. public schools. They affected students of color regardless of specific racial/ethnic groups and were evident in the vast majority of American public schools. Indeed, new evidence for disproportionality across racial/ethnic groups indicated that the problem of racial/ethnic gaps may be worse than has previously been appreciated. In terms of prediction, this study provided strong evidence that the largest gaps in suspension between students of

color and their White counterparts tend to reside in elementary schools that have higher proportions of White students and lower levels of school poverty.

REFERENCES

Adams, J. R., Benshoff, J. M., & Harrington, S. Y. (2007). An examination of referrals to the school counselor by race, gender, and family structure. *Professional School Counseling, 10,* 389–398.

Advancement Project. (2000). Opportunities suspended: The devastating consequences of zero tolerance and school discipline. Cambridge, MA: Civil Rights Project.

Advancement Project. (2005). *Education on lockdown: The schoolhouse to jailhouse track.* Washington, DC: Advancement Project.

Allison, P. (2014). *Listwise deletion: It's not evil.* Retrieved from http:// statisticalhorizons.com/listwise-deletion-its-not-evil.

Allison, P. D. (2001). *Missing data.* Thousand Oaks, CA: Sage.

Alvarez, A. N., Liang, C. H., & Neville, H. A. (2016). *The cost of racism for people of color: Contextualizing experiences of discrimination*. Washington, DC: American Psychological Association.

American Psychological Association Zero Tolerance Task Force. (2006, August 9). *Are zero tolerance policies effective in the schools: An evidentiary review and recommendations?* Retrieved from https://www.apa.org/pubs/info/reports/zero-tolerance-report.pdf

American Psychological Association Zero Tolerance Task Force. (2008). Are zero tolerance policies effective in the schools? An evidentiary review and recommendations. *American Psychologist, 63*(9), 852-862.

Anderson, K. P., & Ritter, G. W. (2017). Disparate use of exclusionary discipline: Evidence on inequities in school discipline from a US state. *Education Policy Analysis Archives*, *25*(49), 1-36.

Annamma, S., Morrison, D., & Jackson, D. (2014). Disproportionality fills in the gaps: Connections between achievement, discipline and special education in the School-to-Prison Pipeline. *Berkeley Review of Education*, *5*(1), 53-87.

Apple, M. W. (1995). *Education and power* (2nd ed.). New York, NY: Routledge.

Arcia, E. (2007). A comparison of elementary/K-8 and middle schools' suspension rates. *Urban Education, 42*(5), 456-469.

Artiles, A. J., Harry, B., Reschly, D. J., & Chinn, P. C. (2002). Over-identification of

children of color in special education: A critical overview. *Multicultural Perspectives, 4,* 3–10.

Artiles, A. A., Kozleski, E. B., Trent, S. C., Osher, D., & Ortiz, A. (2010). Justifying and explaining disproportionality, 1968-2008: A critique of underlying views of culture. *Exceptional Children, 76,* 279–299.

Balfanz, R., Byrnes, V., & Fox, J. H. (2015). Sent home and put off-track: The antecedents, disproportionalities, and consequences of being suspended in the ninth grade. In D. Losen (Ed.), *Closing the school discipline gap: Equitable remedies for excessive exclusion* (pp. 17-30). New York, NY: Teachers College Press.

Bell, D. A. (1973) *Race, racism and American law.* New York, NY: Little, Brown and Company.

Bell, D. A. (1995). Racial realism. In K. Crenshaw, N. Gotanda, G. Peller, & K. Thomas (Eds.), *Critical race theory: The key writings that formed the movement* (pp. 303-312). New York: New Press.

Betts, J. R., Rueben, K. S. & Danenberg, A. (2000). *Equal resources, equal outcomes? The distribution of school resources and student achievement in California*. San Francisco, CA: Public Policy Institute of California.

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science, 15*(10), 674-679.

Blake, J., Butler, B. R., Lewis, C. W., & Darensbourg, A. (2011). Unmasking the inequitable discipline experiences of urban Black girls: Implications for urban educational stakeholders. *Urban Review, 43*(2), 90-106.

Blalock, H. M. (1967). *Toward a theory of minority-group relations*. New York, NY: John Wiley and Sons.

Blumer, H. (1958). Race prejudice as a sense of group position. *The Pacific sociological review, 1*(1), 3-7.

Bobo, L. (1983). Whites' opposition to busing: Symbolic racism or realistic group conflict? *Journal of Personality and Social Psychology*, *45*(6), 1196.

Bollmer, J., Bethel, J., Garrison-Mogren, R., & Brauen, M. (2007). Using the risk ratio to assess racial/ethnic disproportionality in special education at the school-district level. *The Journal of Special Education*, *41*(3), 186-198.

Bollmer, J., Bethel, J., Munk, T., & Bitterman, A. (May, 2014). *Methods for assessing*

racial/ethnic disproportionality in special education: A technical assistance guide (revised). Rockville, MD: IDEA Data Center/Westat.

Bonilla-Silva, E. (2010). *Racism without racists: Color-blind racism and racial inequality in contemporary America* (3rd ed.). Lanham, MD: Rowman and Littlefield.

Bradshaw, C., Mitchell, M., & Leaf, P. (2010). Examining the effects of school-wide positive behavioral interventions and supports on student outcomes: Results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions, 12*(3), 133-148.

Brantlinger, E. (1991). Social class distinctions in adolescents' reports of problems and punishment in school. *Behavioral Disorders, 17*(1), 36–46.

Brooks, K., Schiraldi, V., & Zeidenberg, J. (2001). *School house hype: Two years later.* Washington, DC: Justice Policy Institute, Children's Law Center.

Brown, C. A., & Di Tillio, C. (2013). Discipline disproportionality among Latinx and American Indian students: Expanding the discourse in US research. *Journal of Education and Learning, 2*(4), 47-59.

Butler, B. R. (2011). *The puzzle of discipline: An examination of African American disproportionality in school discipline and student performance*. (Doctoral dissertation). Texas A&M University, College Station, TX.

Butler, B. R., Robinson, D., & Walton, C. W. (2014). A perfect storm: How pose, perception, and threat converge to perpetuate discriminatory discipline practices for Black male students. In C. W. Lewis & J. L. Moore (Eds.), *African American male students in preK-12 schools: Informing research, policy, and practice* (pp. 151-175). Bingley, UK: Emerald Group Publishing Limited.

Bryan, J., Day-Vines, N. L., Griffin, D., & Moore-Thomas, C. (2012). The disproportionality dilemma: Patterns of teacher referrals to school counselors for disruptive behavior. *Journal of Counseling & Development, 90*(2), 177-190.

Cameron, M., & Sheppard, S. M. (2006). School discipline and social work practice: Application of research and theory to intervention. *Children & Schools, 28*(1), 15-22.

Casella, R. (2003). Punishing dangerousness through preventive detention: Illustrating the institutional link between school and prison. In J. Wald & D. J. Losen (Eds.), *New directions for youth development: Deconstructing the school-to-prison pipeline* (55-70). San Francisco, CA: Jossey-Bass.

Casey, P., Warren, R., Cheesman, F., & Elek, J. (2012). *Helping courts address implicit bias: Resources for education.* Williamsburg, VA: National Center for State Courts.

Chamlin, M. B. (1989). A macro social analysis of change in police force size, 1972–1982. *The Sociological Quarterly, 30*(4), 615-624.

Children's Defense Fund. (1975). *School suspensions: Are they helping children?* Cambridge, MA: Washington Research Project.

Christle, C., Nelson, C.M., & Jolivette, K. (2004). School characteristics related to the use of suspension. *Education and Treatment of Children, 27*(4), 509-529.

Colby, S. L. & Ortman, J. M. (2014). *Projections of the size and composition of the U.S. population: 2014 to 2060, current population reports*. Washington, DC: U.S. Census Bureau.

Costello, B. Wachtel, J., & Wachtel, T. (2010). Restorative circles in schools: Building community and enhancing learning. A practical guide for educators. Bethlehem, PA: International Institute for Restorative Practices (IIRP).

Cousineau, D., & Chartier, S. (2010). Outliers detection and treatment: A review. *International Journal of Psychological Research, 3*(1), 58-67.

Delgado, R. & Stefancic, J. (Eds.) (2000) *Critical race theory: The cutting edge*. Philadelphia, PA: Temple University Press.

Delpit, L.D. (2006). *Other people's children: Cultural conflict in the classroom*. New York: New Press.

DeNavas-Walt, C., & Proctor, B. D. (2015). *Income and poverty in the United States: 2014*. Washington, DC: U.S. Government Printing Office.

DeVoe, J. F., Peter, K., Noonan, M., Snyder. T.D., & Baum, K. (2005) *Indicators of school crime and safety 2005*. Washington, DC: US Departments of Education and Justice.

Downey, D., & Pribesh, S. (2004). When race matters: Teachers' evaluations of students' classroom behavior. *Sociology of Education, 77*(4), 267-282.

Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davies, P. G. (2004). Seeing Black: Race, crime, and visual processing. *Journal of Personality and Social Psychology, 87*(6), 876–893.

Edelman, M. W. (2007). The cradle to prison pipeline: An American health crisis. *Preventing chronic disease, 4*(3), 1-2.

Eitle, T. M., & Eitle, D. J. (2004). Inequality, segregation, and the overrepresentation of African Americans in school suspensions. *Sociological Perspectives*, *47*(3), 269-287.

Every Student Succeeds Act, Pub. L. No. 114-95, 129 Stat. 1802 (2015).

Fabelo, T., Thompson, M. D., Plotkin, M. D., Carmichael, D., Marchbanks, M. P., & Booth, E. A. (2011). *Breaking schools' rules: A statewide study of how school discipline relates to students' success and juvenile justice involvement.* New York, NY: Council of State Governments Justice Center.

Fenning, P., Wilczynski, J., & Parraga, M. (2000). A comparative analysis of existing secondary school discipline policies: Implications for improving practice and school safety. In McClafferty, K., C.A. Torres, & T.R. Mitchell (Eds.), *Challenges of urban education: Sociological perspectives for the next century* (237-252). Albany, NY: State University of New York Press.

Fenning, P., & Rose, J. (2007). Overrepresentation of African American students in exclusionary discipline: The role of school policy. *Urban Education*, *42*(6), 536-559.

Fenning, P., Pigott, T., Engler, E., Bradshaw, K., Gamboney, E., Grunewald, S., & McGrath-Kato, M. (2013). A mixed methods approach examining disproportionality in school discipline. *Closing the school discipline gap: Research to practice, Washington, DC*.

Ferguson, A. A. (2001). *Bad boys: Public schools in the making of Black masculinity.* Ann Arbor, MI: University of Michigan Press.

Finn, J. D., Fish, R. M., & Scott, L. A. (2008). Educational sequelae of high school misbehavior. *The Journal of Educational Research*, *101*(5), 259-274.

Finn, J. D., & Servoss, T. J. (2013). Misbehavior, suspensions, and security measures in high school: Racial/ethnic and gender differences. In D. Losen (Ed.), *Closing the school discipline gap: Research for policymakers* (pp. 44–58). New York, NY: Teachers College Press.

Ford, C. L., & Airhihenbuwa, C. O. (2010). Critical race theory, race equity, and public health: toward antiracism praxis. *American Journal of Public Health*, *100* (supplement 1), S30-S35.

Freeman, K. J., & Steidl, C. R. (2016). Distribution, composition and exclusion: How School segregation impacts racist disciplinary patterns. *Race and Social Problems*, *8*(2), 171-185.

Garson, D. G. (2015). *Missing values analysis and data imputation.* Asheboro, NC: Statistical Publishing Associates.

Gay, G. (2006). Connections between classroom management and culturally responsive teaching. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management*. (pp. 343-372). Mahwah, NJ: Lawrence Erlbaum.

Giroux, H. A. (2003). Racial injustice and disposable youth in the age of zero tolerance. *International Journal of Qualitative Studies in Education, 16*(4), 553–65.

Glock, S., & Karbach, J. (2015). Preservice teachers' implicit attitudes toward racial minority students: Evidence from three implicit measures. *Studies in Educational Evaluation, 45,* 55–61.

Glock, S. (2016). Stop talking out of turn: The influence of students' gender and ethnicity on preservice teachers' intervention strategies for student misbehavior. *Teaching and Teacher Education, 56,* 106-114.

Gottfredson, G. D., Gottfredson, D. C., Czeh, E. R., Cantor, D., Crosse, S. B. & Hantman, I. (2000) *A national study of delinquency prevention in school: Final report.* Ellicott City, MD: Gottfredson Associates.

Green, E. L. (2017, June 16). Education dept. says it will scale back civil rights investigations. *The New York Times*. Retrieved from https://www.nytimes.com/2017/06/16/us/politics/education-department-civil-rights-betsy-devos.html

Gregory, A., & Weinstein, R.S. (2008). The discipline gap and African Americans: Defiance or cooperation in the high school classroom. *Journal of School Psychology, 46*(4), 455-475.

Gregory, A., & Thompson, A. R. (2010). African American high school students and variability in behavior across classrooms. *Journal of Community Psychology, 38*(3), 386-402.

Gregory, A., Skiba, R. J. & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher*, *39*(1), 59-68.

Gregory, A., Bell, J., & Pollock, M. (2014). How educators can eradicate disparities in school discipline: A briefing paper on school-based interventions. *Discipline Disparities Series: Interventions, March. The Equity Project at Indiana University, Center for Evaluation and Education Policy*.

Gregory, A., Clawson, K., Davis, A., & Gerewitz, J. (2016). The promise of restorative

practices to transform teacher-student relationships and achieve equity in school discipline. *Journal of Educational and Psychological Consultation*, *26*(4), 325-353.

Gregory, A., Skiba, R. J., & Mediratta, K. (2017). Eliminating disparities in school discipline: A framework for intervention. *Review of Research in Education*, *41*(1), 253-278.

Grothaus, T. (2013). School counselors serving students with disruptive behavior disorders. *Professional School Counseling, 16*(4), 245-255.

Hamre, B. K., & Pianta, R. C. (2006). Student-teacher relationships. In G. G. Bear & K. M. Minke (Eds.), *Children's needs III: Development, prevention, and intervention* (pp. 59–71). Bethesda, MD: National Association of School Psychologists.

Harris, A. P. (2012). *Critical race theory*. Davis, CA: University of California Davis. Heller, M. (2007). Bilingualism as ideology and practice. In M. Heller (Ed.), Bilingualism: a social approach (pp. 1–24). New York: Palgrave Macmillan.

Hindelang, M. J., Hirschi, T., & Weis, J. G. (1979). Correlates of delinquency: The illusion of discrepancy between self-report and official measures. *American Sociological Review, 44*(6), 995–1014.

Hinojosa, M. S. (2008). Black-white differences in school suspension: Effect on student beliefs about teachers. *Sociological Spectrum, 28*(2), 175–193.

Hinojosa, M. S., & Moras, A. (2009). Challenging colorblind education: A descriptive analysis of teacher racial attitudes. *Research and Practice in Social Sciences, 4*(2), 27-45.

Hirschfield, P. J. (2008). Preparing for prison? The criminalization of school discipline in the USA. *Theoretical Criminology*, *12*(1), 79-101.

Hirschfield, P. J., & Celinska, K. (2011). Beyond fear: Sociological perspectives on the criminalization of school discipline. *Sociology Compass, 5*(1), 1-12.

Horner, R., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A. W., & Esperanza, J. (2009). A randomized, wait-list controlled effectiveness trial assessing school-wide positive behavior support in elementary schools. *Journal of Positive Behavioral Interventions*, *11*(3), 133–144.

Horner, S. B., Fireman, G. D., & Wang, E. W. (2010). The relation of student behavior, peer status, race, and gender to decisions about school discipline using CHAID decision trees and regression modeling. *Journal of School Psychology, 48*(2), 135-161.

Hosp, J. L., & Reschly, D. J. (2003). Referral rates for intervention or assessment: A meta-analysis of racial differences. *The Journal of Special Education, 37*, 67–80.

Howard, T. C. (2003). Culturally relevant pedagogy: Ingredients for critical teacher reflection. *Theory into practice*, *42*(3), 195-202.

Hsin, A., & Xie, Y. (2014). Explaining Asian Americans' academic advantage over whites. *Proceedings of the National Academy of Sciences*, *111*(23), 8416-8421.

Huang, F. L. (2016). Do Black students misbehave more? Investigating the differential involvement hypothesis and out-of-school suspensions. *The Journal of Educational Research*, 1-11.

Jacob, B. A. (2007). The challenges of staffing urban schools with effective teachers. *The Future of Children, 17*(1), 129-153.

Jacobs, D., & Kleban, R. (2003). Political institutions, minorities, and punishment: a pooled cross-national analysis of imprisonment rates. *Social Forces*, *82*(2), 725-755.

Kang, J., Bennet, M., Carbado, D., Casey, P., Dasgupta, N., Faigman, D., Godsil, R., Greenwald, A. G., Levinson, J. & Mnookin, J. (2012). Implicit bias in the courtroom. *UCLA Law Review*, *59*(1), 1124-1186.

Kewel Ramani, A., Gilbertson, L., Fox, M., & Provasnik, S. (2007). *Status and trends in the education of racial and ethnic minorities* (NCES 2007-039). National Center for Educational Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Kinder, D. R., & Sears, D. O. (1981). Prejudice and politics: Symbolic racism versus racial threats to the good life. *Journal of Personality and Social Psychology*, *40*(3), 414.

Kluckhohn, C. (1962). *Culture and behavior*. New York, NY: Free Press of Glencoe.

Krezmien, M. P., Leone, P. E., & Achilles, G. M. (2006). Suspension, race, and disability: Analysis of statewide practices and reporting. *Journal of Emotional and Behavioral Disorders, 14*(4), 217-226.

Kupchik, A., & Ward, G. K. (2011). Reproducing social inequality through school security: Effects of race and class on school security measures. Proceedings from: *The 106th Annual Meeting of the American Sociological Association*. Las Vegas, NV.

Kyle, P. B., & DeVoss, J. A. (2015). The professional school counselor's role as change agent in school. *School Counseling Research & Practice*, *3*, 7-16.

Ladson-Billings, G., & Tate, W. F. (1995). Toward a critical race theory of education. *Teachers College Record*, *97*(1), 47-68.

Ladson-Billings, G. (1998). Just what is critical race theory and what's it doing in a nice field like education? *International Journal of Qualitative Studies in Education*, *11*(1), 7-24.

Lankford, H., Loeb, S. & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Education Evaluation and Policy Analysis, 24*(1), 37-62.

Lassen, S. R., Steele, M. M., & Sailor, W. (2006). The relationship of school-wide positive behavior support to academic achievement in an urban middle school. *Psychology in the Schools, 43*(6), 701-712.

Lemert, C. (Ed.). (2010). *Social theory: The multicultural and classic readings (4th Edition)*. Boulder, CO: Westview Press.

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin,125*(2), 255–275.

Lewis, C.W., Butler, B.R., Bonner, F.A., & Joubert, M. (2010). African American male discipline patterns and school district responses resulting impact on academic achievement: Implications for urban educators and policy makers. *Journal of African American Males in Education*, *1*(1), 1-19.

Losen, D., & Skiba, R. (2010). *Suspended education: Urban middle schools in crisis.* Montgomery, AL: Southern Poverty Law Center.

Losen, D. J., & Gillespie, J. (2012). *Opportunities suspended: The disparate impact of disciplinary exclusion from school.* Los Angeles, CA: The Civil Right Project, The Center for Civil Rights Remedies.

Losen, D. J., & Martinez, T. E. (2013). *Out of school & off track: The overuse of suspensions in American middle and high schools.* Los Angeles, CA: The Center of Civil Rights Remedies.

McLoughlin, C. S., & Noltemeyer, A. L. (2010). Research into factors contributing to discipline use and disproportionality in major urban schools. Current Issues in Education, 13(2).

Mann, H. (1849). Twelfth annual report of Horace Mann as Secretary of Massachusetts State Board of Education. Boston, MA: Dutton and Wentworth.

Maryland Disability Law Center. (2015). *The school discipline process: A handbook for*

Maryland families and professionals. Baltimore, MD: Maryland Disability Law Center.

McCarthy, J. D., & Hoge, D. R. (1987). The social construction of school punishment: Racial disadvantage out of universalistic process. *Social Forces, 65*(4), 1101–1120.

McConahay, J. B. (1982). Self-interest versus racial attitudes as correlates of anti-busing attitudes in Louisville: Is it the buses or the Blacks? *The Journal of Politics, 44*(3), 692-720.

McFadden, A. C., Marsh, G. E., Price, B. J., & Hwang, Y. (1992). A study of race and gender bias in the punishment of school children. *Education and Treatment of Children, 15*(2), 140-146.

Minneapolis Public Schools. (2014, November). *Positive schoolwide engagement plan: Behavior standards discipline policy for Minneapolis Public Schools (MPS)*. Minneapolis, MN: Minneapolis Public Schools. Retrieved from http://www.mpls.k12.mn.us/uploads/behavior_standards_for_web.pdf

Monroe, C. R. (2013). Discipline and diversity in the suburban US South. *Race Ethnicity and Education, 16*(2), 182-202.

Morrison, G. M., & Skiba, R. J. (2001). Predicting violence from school misbehavior: Promises and perils. *Psychology in the Schools, 38*(2), 173–184.

Mosher, C. (2001). Predicting drug arrest rates: Conflict and social disorganization perspectives. *Crime & Delinquency. 47*(1), 84–104.

Mukuria, G. (2002). Disciplinary challenges: How do principals address this dilemma? *Urban Education, 37*(3), 432-452.

Murray, C., & Herrnstein, R. J. (1994). *The bell curve: Intelligence and class structure in American life.* New York: Free Press.

National Center for Educational Statistics. (2017, March). *The condition of education: English language learners in public schools.* Retrieved from https://nces.ed.gov/programs/coe/indicator_cgf.asp

Neal, L. I., McCray, A. D., Webb-Johnson, G., & Bridgest, S. T. (2003). The effects of African American movement styles on teachers' perceptions and reactions. *Journal of Special Education, 37*(1), 49–57.

Neuberg, S. L., & Cottrell, C. A. (2002). Intergroup emotions: A biocultural approach. In

D. Mackie & E. Smith (Eds.), *From prejudice to intergroup emotions: Differentiated reactions to social groups* (pp. 265-283). New York: Psychology Press.

Nichols, J. D. (2004). An exploration of discipline and suspension data. *The Journal of Negro Education, 73*(4), 408–423.

Nicholson-Crotty, S., Birchmeier, Z., & Valentine, D. (2009). Exploring the impact of school discipline on racial disproportion in the juvenile justice system. *Social Science Quarterly, 90*(4), 1003-1018.

Noltemeyer, A. L., Ward, R. M., & McLoughlin, C. (2015). Relationship between school suspension and student outcomes: A meta-analysis. *School Psychology Review, 44*(2), 224-240.

O'Brien, R. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity, 41*(5), 673-690.

Okonofua, J. A., & Eberhardt, J. L. (2015). Two strikes: Race and the disciplining of young students. *Psychological Science, 26*(5), 617-624.

Oregon Department of Public Education. (2014, August). *Finding significant disproportionality*. Retrieved from http://www.ode.state.or.us/search/page/?=3149

Orfield, G. & Frankenberg, E. (2014). Brown at 60: Great progress, a long retreat and an uncertain future. *The Civil Rights Project/Proyecto Derechos Civiles.* The University of California Los Angeles.

Parker, L., & Lynn, M. (2002). What's race got to do with it? Critical race theory's conflicts with and connections to qualitative research methodology and epistemology. *Qualitative Inquiry, 8*(1), 7-22.

Parrish, T. (2002). Racial disparities in the identification, funding, and provision of special education. In D. J. Losen & G. Orfield (Eds.), *Racial inequity in special education* (pp. 15–37). Cambridge, MA: Harvard Civil Rights Project.

Payne, A. A., & Welch, K. (2010). Modeling the effects of racial threat on punitive and restorative school discipline practices. *Criminology, 48*(4), 1019-1062.

Peguero, A. A., & Shekarkhar, Z. (2011). Latino/a student misbehavior and school punishment. *Latinx Journal of Behavioral Sciences, 33*(1), 54-70.

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

Pew Research Center. (2015). *Multiracial in America: Proud, diverse and growing in numbers.* Washington, D.C.: Pew Research Center.

Pit-ten Cate, I. M., Krolak-Schwerdt, S., & Glock, S. (2016). Accuracy of teachers' tracking decisions: Short-and long-term effects of accountability. *European Journal of Psychology of Education, 31*(2), 225-243.

Poe-Yamagata, E., & Jones, M. (2000). *And justice for some.* Washington, DC: Building Blocks for Youth.

Public Counsel Law Center. (2012). *Fix school discipline: How we can fix school discipline toolkit for educators.* Los Angeles, CA: Public Counsel Law Center.

Raffaele Mendez, L. M., Knoff, H. M., & Ferron, J. M. (2002). School demographic variables and out-of-school suspension rates: A quantitative and qualitative analysis of a large, ethnically diverse school district. *Psychology in the Schools*, *39*(3), 259-277.

Raffaele Mendez, L. M. (2003). Predictors of suspension and negative school outcomes: A longitudinal investigation. *New Directions for Youth Development, 99*(1), 17-33.

Raffaele Mendez, L. M., & Knoff, H. M. (2003). Who gets suspended from school and why: A demographic analysis of schools and disciplinary infractions in a large school district. *Education and Treatment of Children*, *26*(1), 30-51.

Rausch, M. K., & Skiba, R. J. (2006). Exclusion is not the only alternative: The children left behind project. In A. Reyes (Ed.), Discipline, achievement, and race: Is zero tolerance the answer? (pp. 105–126). Lanham, MD: Rowman & Littlefield Publishing Group.

Reschly, D. J. (1997). *Disproportionate minority representation in general and special education: Patterns, issues, and alternatives.* Des Moines: Iowa Department of Education.

Reskin, B. (2005). Unconsciousness raising. *Regional Review, 14*(3), 32–37.

Reyes, A. H. (2006). *Discipline, achievement, and race: Is zero tolerance the answer?* Lanham, MD: Rowman and Littlefield.

Riek, B. M., Mania, E. W., & Gaertner, S. L. (2006). Intergroup threat and outgroup attitudes: A meta-analytic review. *Personality and Social Psychology Review, 10*(4), 336-353.

Rocha, R., & Hawes, D. (2009). Racial diversity, representative bureaucracy, and

equity in multicultural districts. Social Science Quarterly, 90(2), 326–344.

Rocque, M. (2010). Office discipline and student behavior: Does race matter? *American Journal of Education*, *116*(4), 557-581.

Rocque, M., & Paternoster, R. (2011). Understanding the antecedents of the "school-to-jail" link: The relationship between race and school discipline. *Journal of Criminal Law and Criminology, 101*(2), 633-666.

Rothery, P. (1988). A cautionary note on data transformation: bias in back-transformed means. *Bird Study*, *35*(3), 219-221.

Schiff, M. (2013, January). *Dignity, disparity and desistance: Effective restorative justice strategies to plug the "school-to-prison pipeline."* Paper presented at the Closing the School Discipline Gap: Research to Practice conference, Washington, DC.

Sears, D. 0. (1988). Symbolic racism. In P. A. Katz & D. A. Taylor (Eds.), *Eliminating racism: Profiles in controversy* (pp. 53-84). New York: Plenum Press.

Shaw, S. R., & Braden, J. P. (1990). Race and gender bias in the administration of corporal punishment. *School Psychology Review*, *19*(3), 378-383.

Sherif, M., & Sherif, C. W. (1969). In group and intergroup relations: Experimental analysis. In M. Sherif & C. W. Sherif (Eds.), *Social psychology* (pp. 221-266). New York: Harper & Row.

Skiba, R. J., Peterson, R. L., & Williams, T. (1997). Office referrals and suspension: Disciplinary intervention in middle schools. *Education and treatment of children*, *20*(3), 295-315.

Skiba, R., & Peterson, R. (1999). The dark side of zero tolerance: Can punishment lead to safe schools? *The Phi Delta Kappan*, *60*(5), 372-376, 381-382.

Skiba, R. (2000). Zero tolerance, zero evidence: An analysis of school disciplinary practice. *Indiana Education Policy Center*. Policy research report #SRS2. Bloomington, IN: Indiana University, Indiana Education Policy Center.

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

Skiba, R. J., & Rausch, M. K. (2006). Zero tolerance, suspension, and expulsion: Questions of equity and effectiveness. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice, and contemporary issues* (pp. 1063–1089). Mahwah, NJ: Lawrence Erlbaum Associates.

Skiba, R. J., Eckes, S. E. & Brown, K. (2010) African American disproportionality in school discipline: The divide between best evidence and legal remedy. *New York Law School Law Review, 54,* 1071-1088.

Skiba, R. J., Horner, R. H., Chung, C., Rausch, M. K., May, S., & 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., Arredondo, M. I., & Rausch, M. K. (2014). *New and developing research on disparities in discipline*. Bloomington, IN: The Equity Project at Indiana University.

Skiba, R. J., Chung, C. G., Trachok, M., Baker, T., Sheya, A., & Hughes, R. (2014). Where should we intervene? Contributions of behavior, student, and school characteristics to out-of-school suspension. In D. Losen (Ed.), *Closing the school discipline gap: Equitable remedies for excessive exclusion* (pp. 132-146). New York, NY: Teachers College Press.

Sleeter, C. E. (2012). Confronting the marginalization of culturally responsive pedagogy. *Urban Education, 47*(3), 562-584.

Smith, R., & Lowery, L. (2017). Students can't wait: Promoting equity and improvement through ESSA. *State Education Standard, 17*(3), 7-10.

Snyder, T.D., de Brey, C., and Dillow, S.A. (2016). *Digest of education statistics 2015* (NCES 2016-014). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Snyder, T. D., & Musu-Gillette, L. (2015, April). Free or reduced price lunch: A proxy for poverty? Washington, DC: National Center for Educational Statistics. Retrieved from https://nces.ed.gov/blogs/nces/post/free-or-reduced-price-lunch-a-proxy-for-poverty

Sparks, S. (2014, August). Popular child-poverty measure gets another look. *Education Week, 34*(1). Retrieved from http://www.edweek.org/ew/articles/2014/08/20/01povertymeasures.h34.html

Solarzano, D. G., & Bernal, D. D. (2001). Examining transformational resistance through critical race and LatCrit theory framework: Chicana and Chicano students in urban context. *Urban Education, 36*(3), 308-342.

Saporu, D. F. (2012). *Suspended opportunities? A multi-level analysis of the role of school climate and composition in shaping racial differences in school punishment* (Doctoral dissertation). The Ohio State University, Columbus, OH.

St. George, D. (2011, June 1). More schools are rethinking zero tolerance. *The Washington Post.* Retrieved from https://www.washingtonpost.com/local/education/more-schools-are-rethinking-zero-tolerance/2011/05/26/AGSIKmGH_story.html?utm_term=.e7dfb7e2aed6

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American psychologist*, *52*(6), 613-629.

Stephan, W. G., & Stephan, C. W. (1996). Predicting prejudice. *International Journal of Intercultural Relations*, *20*(3), 409-426.

Sullivan, A. L., Klingbeil, D. A., & Van Norman, E. R. (2013). Beyond behavior: Multilevel analysis of the influence of socio-demographics and school characteristics on students' risk of suspension. *School Psychology Review, 42*(1), 99–114.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson Education, Inc.

Thornton, C. H., & Trent, W. T. (1988). School desegregation and suspension in East Baton Rouge Parish: A preliminary report. *The Journal of Negro Education*, *57*(4), 482-501.

Townsend, B. L. (2000). The disproportionate discipline of African American learners: Suspensions and expulsions. *Exceptional Children*, *66*(3), 381-291.

Trachok, M. (2015). *Teacher interview as a methodology for assessing implicit bias: A qualitative investigation* (Doctoral dissertation). Indiana University, Bloomington, IN.

Triplett, N. P., Allen, A., & Lewis, C. W. (2014). Zero tolerance, school shootings, and the post-Brown quest for equity in discipline policy: An examination of how urban minorities are punished for White suburban violence. *The Journal of Negro Education*, *83*(3), 352-370.

Triplett, N. P., Bryant, A. C., Brown, K. E., Steele, A. S. L., Ardrey, T., Allen, A. & Lewis, C. W. (2016). Discipline disproportionality, student achievement and the Every Student Succeeds Act (ESSA). *Educational Law and Policy Review, 3*(1), 209-240.

Trochim, W. M., & Donnelly, J. P. (2006). The research methods knowledge base (3rd ed.). Cincinnati, OH: Atomic Dog.

U.S. Census Bureau. (2015). Projections of the size and composition of the U.S. population: 2014 to 2060: Estimates and projections. Washington, DC. U.S. Census Bureau.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (n.d.) Identification of rural locales. Retrieved from http://nces.ed.gov/ccd/ rural_locales.asp

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (1996). *Urban schools: The challenge of location and poverty*. (NCES 96-184). Washington, DC: U.S. Department of Education. Retrieved from http://nces.ed.gov/pubs/96184all.pdf

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2007). *Status and trends in the education of racial and ethnic minorities*. (NCES 2007-039). Washington, DC: U.S. Department of Education. Retrieved rom http://nces.ed.gov/pubs2007/minoritytrends/tables/table_7_1.asp?referrer=report

U.S. Department of Education. (2009a). *Schools and staffing survey*. (NCES 2009-324). Washington, DC: U.S. Department of Education. Retrieved from http://nces.ed.gov/pubs2009/2009324/tables/sass0708_2009324_t12n_02.asp

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2009b). *Schools and staffing survey*. (NCES 2009-324). Washington, DC: U.S. Department of Education. Retrieved from http://nces.ed.gov/pubs2009/2009324/tables/sass0708_2009324_t12n_02.asp

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2009c). *Schools and staffing survey*. Washington, DC: U.S. Department of Education. Retrieved from http://nces.ed.gov/surveys/sass/tables/sass0708_2013027_s1n_02.asp

U.S. Department of Education. (2010a). *Status and trends in the education of racial and ethnic minorities*. (NCES 2010-015). Washington, DC. U.S. Department of Education. Retrieved from http://nces.ed.gov/pubs2010/2010015/tables/table_7_2.asp

U.S. Department of Education. (2010b). *A blueprint for reform: The reauthorization of the Elementary and Secondary Education Act*. Retrieved from https://www2.ed.gov/policy/elsec/leg/blueprint/blueprint.pdf

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2012a). *Schools and staffing survey*. Retrieved from https://nces.ed.gov/surveys/sass/tables/sass1112_2013314_t1s_001.asp

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2012b). *Schools and staffing survey*. Retrieved from https://nces.ed.gov/surveys/sass/tables/sass1112_490_a1n.asp

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2013). *Digest of educational statistics.* Washington DC: U.S. Department of Education. Retrieved from http://nces.ed.gov/programs/digest/d13/tables/dt13_214.40.asp

U.S. Department of Justice & U.S. Department of Education. (2014). *School Climate and Discipline*. Washington, DC. U.S. Department of Education. Retrieved from http://www2.ed.gov/policy/gen/guid/school-discipline/index.html

U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), (2015). Local education agency (school district) finance survey (F33), 2011-12. Retrieved from: https://nces.ed.gov/programs/digest/d14/tables/dt14_235.40.asp

U.S. Department of Education. (2016a). *Oklahoma City public schools, U.S. education department reach settlement to address disproportionate discipline of Black students.* Washington DC: U.S. Department of Education. Retrieved from http://www.ed.gov/news/press-releases/oklahoma-city-public-schools-us-education-department-reach-settlement-address-disproportionate-discipline-black-students

U. S. Department of Education. (2016b). *2013-2014 Civil Rights Data Collection: A first look.* Washington, DC: U.S. Department of Education Office for Civil Rights

Van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M., & Holland, R. W. (2010). The implicit prejudiced attitudes of teacher's relations to teacher expectations and the ethnic achievement gap. *American Educational Research Journal*, *47*(2), 497-527.

Vavrus, F, & Cole, K. (2002). "I didn't do nothin'": The discursive construction of school suspension. *The Urban Review, 34*(2), 87–111.

Wald, J., & Losen, D. (2003). Defining and redirecting a school-to-prison pipeline. *New Directions in Youth Development, 99,* 9–15.

Wallace, J. M. Jr., Goodkind, S., Wallace, C. M., & Bachman, J. G. (2008). Racial, ethnic, and gender differences in school discipline among U. S. high school students: 1991-2005. *Negro Educational Review, 59*(1-2), 47-62.

Warren, J. S., Edmonson, H. M., Griggs, P., Lassen, S. R., McCart, A., Turnbull, A. & Sailor, W. (2003) Urban applications of school-wide positive behavior support: Critical issues and lessons learned. *Journal of Positive Behavioral Interventions, 5*(2), 80-92.

Watts, I. E., & Erevelles, N. (2004) These deadly times: Reconceptualizing school violence by using critical race theory and disability studies. *American Educational Research Journal, 41*(2), 271-299.

Welch, K. & Payne, A. A. (2010). Racial threat and punitive school discipline. *Social Problems, 57*(1), 25-48.

Wiley, T. G., & Lukes, M. (1996). English-only and standard English ideologies in the US. *TESOLl Quarterly, 30*(3), 511-535.

Williams, N. A. (2015). *Are the racial disparities in school discipline the result of or a function of systemic racism mediated by educators' dispositions?* (Doctoral dissertation). Indiana University, Bloomington, IN.

Wilson, J. Q., & Herrnstein, R. J. (1985). *Crime and human nature*. New York: Oxford.

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

Zarate, M. A., Garcia, B., Garza, A. A., & Hitlan, R. T. (2004). Cultural threat and perceived realistic group conflict as dual predictors of prejudice. *Journal of Experimental Social Psychology, 40*(1), 99-105.

APPENDIX A: SAMPLE DETAILS BY RACIAL/ETHNIC GROUP

Sample Details by Racial/Ethnic Group

	Mean	Range	Min	Max
AI/AN (n=4017)				
AI/AN Enrollment	70.53	1458	10	1468
White Enrollment	399.90	3288	10	3298
Asian (n=7786)				
Asian Enrollment	111.90	1992	10	2002
White Enrollment	500.52	3705	10	3715
Black (n=29196)				
Black Enrollment	153.04	2586	10	2596
White Enrollment	374.07	3288	10	3298
Latinx (n=29526)				
Latinx Enrollment	223.23	4254	10	4264
White Enrollment	366.56	3288	10	3298
NH/PI (n=1089)				
Asian Enrollment	55.70	1206	10	1216
White Enrollment	394.57	2349	10	2359
TomR (n=17853)				
TomR Enrollment	37.63	753	10	763
White Enrollment	423.50	3102	10	3112

Note: AI/AN = American Indian/Alaska Native, NH/PI = Native Hawaiian/ Pacific/Islander, TomR = Two or more races

APPENDIX B: SAMPLE SIZE BY PREDOMINANTLY-WHITE STATUS

Sample Size by Predominantly-White Status

	AI/AN	Asian	Black	Latinx	NH/PI	TomR
Not Predominantly-White	1820	4361	14802	14994	838	7938
Predominantly-White	2197	3425	14394	14532	251	9915
Total	4017	7786	29196	29526	1089	17853

Note: AI/AN = American Indian/Alaska Native, NH/PI = Native Hawaiian/ Pacific Islander, TomR = Two or more races

APPENDIX C: SCHOOL LEVEL RELATIVE RISK RATIO OF SUSPENSION BY THE PROPORTION OF WHITE STUDENTS (QUARTILES)

School Level Relative Risk Ratio of Suspension by the Proportion of White Students (Quartiles)

Quartile	AI/AN	Asian	Black	Latinx	NH/PI	TomR
1 (0- 25% White)	2.58	0.85	1.92	0.82	1.60	2.47
2 (25-50% White)	3.21	1.35	3.19	1.49	2.43	3.41
3 (50-75% White)	3.92	1.89	4.21	2.16	3.48	4.02
4 (75-100% White)	4.94	3.01	5.84	3.68	5.38	5.42
Total (Mean RRR)	3.66	1.77	3.79	2.03	3.27	3.83

Note: AI/AN = American Indian/Alaska Native, NH/PI = Native Hawaiian/ Pacific Islander, TomR = Two or more races, RRR = relative risk ratio

APPENDIX D: THE PROPORTION OF WHITE STUDENTS BY RACIAL/ETHNIC GROUP SAMPLE

The Proportion of White Students by Racial/Ethnic Group Section

	Full Sample	AI/AN	Asian	Black	Latinx	NH/PI	TomR
N	89467	4017	7786	29196	29526	1089	17853
Mean	.479	.501	.439	.474	.474	.312	.518
Median	.494	.526	.446	.489	.487	.273	.540
Percentiles							
25	.246	.304	.210	.238	.231	.112	.308
50	.494	.526	.446	.489	.487	.273	.540
75	.705	.703	.652	.702	.708	.482	.735

Note: AI/AN = American Indian/Alaska Native, NH/PI = Native Hawaiian/ Pacific Islander, TomR = Two or more races