Theses and Dissertations

2017

Implicit Bias in School Disciplinary Decisions

Gina Laura Gullo
Lehigh University

Follow this and additional works at: http://preserve.lehigh.edu/etd

Part of the Educational Leadership Commons

Implicit Bias in School Disciplinary Decisions

by

Gina Laura Gullo

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Doctor of Education

in

Educational Leadership

Lehigh University

April 13, 2017

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Education.

_____
Date

_____
Dr. Floyd Beachum
Dissertation Chair

_____
Accepted Date

Committee Members:

_____
Dr. Jill Sperandio

_____
Dr. Qiong Fu

_____
Dr. Leticia Smith-Evans Haynes

iii

ACKNOWLEDGEMENTS

me why I was working so hard, while Dr. Zirkel never accepted anything less than perfection from me and made sure I knew it. Outside of the College of Ed, Kathleen Hutnick's mentorship has been beyond words. She is always there for me and everywhere all at once and I will never understand how, but will also never forget how much. Lehigh has a big place in my heart, and I know that my mentors here were and are a big part of why.

Friends are not easy to come by or keep when living the graduate student life, but I am lucky enough to have quite a few who led my path here. Linda Mayger was one of the first people who helped me to realize that I was a researcher and not just a student of research. Our conversations about education are one of my favorite things, and I can't imagine thinking through my dissertation without her feedback. I know that big things will come from those continued conversations…let the writing begin! Lauryn Woodman-Artis has been crucial to my sanity and progress over the last year. Although some of her help was through her job, most of it was through friendship. Chats, venting, wine, and not talking about my research were as contributory as the education to beginning to understand what kind of researcher I am and what that means. Lauryn made sure I had something to laugh about and someone to talk to just by being there to support me. Both Linda and Lauryn inspired me, comforted me, listened to me, and much more; I am beyond grateful.

Finally, my family is the foundation of my success. My parents supported me in their own non-traditional ways with questions, challenges, and confusion. Growing up with less provided me with the opportunity to empathize with those discriminated against every day. The real resources of love and care that are plentiful from my parents and

siblings led me to grow into the woman I am now. My mother raised me to care for everyone, my father taught me to challenge beliefs, my stepmother showed me how to persevere, and my in-laws provided an endless supply of encouragement. As a parent myself, my son created an extra challenge towards completion, but he's brought so much joy and the purest kind of love into my heart every day. Above all, my husband supported me and encouraged me every day and every step of the way. Dominic Gullo is a man without compare with his unending patience, love, and everything. Without him, I would not have a dissertation and certainly would not be the woman I am today. Dominic, my love, brings happiness and joy to my every day no matter how I wake up and how hard the path for that day. I'm so lucky to have him.

I am here and this paper is written because of so many people. A dissertation is clearly not a project completed alone, and I dedicate this to all of those who helped me along the journey. Thank you.

TABLE OF CONTENTS

CHAPTER 1

CHAPTER 2

CHAPTER 4

    Results 78

CHAPTER 5

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABSTRACT

School discipline attempts to keep students safe, but often disservices students

most in need of education (Children's Defense Fund, 1975; Perry & Morris, 2014; Skiba

& Rausch, 2004). Exclusionary discipline affects students of Color[1] disparately due to

overrepresentation in suspensions and expulsions (Losen & Skiba, 2010). The discipline

gap between White students and students of Color continues to grow—particularly for

Black students (Losen, et al, 2015). Investigations of exclusionary discipline use

suggested school-level factors including administrators' disciplinary preferences

explained more of this phenomenon than other factors and may work through implicit

biases (Skiba, et al, 2014). Implicit biases (attitudes or stereotypes held subconsciously

and unintentionally acted upon) exist for all kinds of preferences but are most concerning

for racial stereotypes with respect to school discipline gaps (Straats, Capatosto, Wright,

& Jackson, 2016). If administrators hold racial implicit biases, one might expect disparate

rates of discipline severity for less favorable groups.

In order to determine whether and to what degree administrators' racial implicit

biases explained discrepant discipline, student discipline data and the Implicit

Associations Test were examined. Administrators' implicit bias scores were related to

the student race to discipline severity relationship and local discipline gaps with

separation by decision types (subjective or objective). For overall and subjective

decisions, implicit bias accounted for differences in the student race to discipline severity

relationship, but only subjective findings remained after demographic and behavioral

---

[1] Color is capitalized throughout this paper when referring to race to mirror the capitalization of White and replicate the capitalization of the more politically acceptable terms: African American and Hispanic. Due to the study aim to better understand bias as it relates to race, people of Color is a more accurate term to describe those individuals who may experience bias due to an increased level of skin pigmentation.

controls. Students of Color experience more severe discipline as a function of administrator implicit bias in subjective discipline decisions. This is the first study to the author's knowledge to demonstrate administrator-level impact of implicit bias on racial discipline discrepancies.

CHAPTER 1

Introduction

Throughout US history the intense interplay between discipline and bias has directed political outcries for equity, equality, and the difficult balance between the two. From Jim Crow laws creating a divide between White Americans and Americans of Color to modern conflicts giving way to the Black Lives Matter movement, bias-based challenges repeatedly acted as a plague on the criminal justice system (Lee, 2013). Derivatives of this implicit bias are present in the US educational system in the form of the discipline gap between White students and students of Color (Straats, Capatosto, Wright, & Jackson, 2016). The school discipline gap demonstrates a ravine between the exclusionary discipline (any discipline that removes a student from the learning environment) rates of these populations of students, with Black and Hispanic students receiving many more incidents of exclusionary discipline than White students. This gap could result from many different issues including implicit bias of those making behavioral consequence decisions (Carter, Skiba, Arredondo, & Pollock, 2014; Kahn, Goff, & Glaser, 2016).

Six in 100 K-12 students in the US received an out of school suspension (OSS) in 2014, but 18% of Black male and 10% of Black female students as opposed to only 5% of White male and 2% of White female students experienced an OSS (CRDC, 2016). This resulted in Black students experiencing suspensions nearly 3.8 times more often than White students. Even Black preschoolers experienced this gap with Black public preschool students representing 19% of enrollment but 47% of students receiving one or more OSS. White public preschool students represented 47% of those enrolled and only

accounted for 28% of students receiving OSS. With students experiencing racial disparities in exclusionary discipline as early as preschool and into elementary and secondary school, the existence of a racial school discipline gap remains undeniable.

The school discipline gap presents a series of concerns for students, communities, and society (Marchbanks, et al., 2015; Skiba, et al., 2013). Receiving just one OSS increases students' risk of dropping out up to 42%, increases student absences from school, and is negatively correlated with academic success (Balfanz, Byrnes, & Fox, 2013; Morris & Perry, 2016). Getting suspended often precedes more serious delinquency and contributes to the School to Prison Pipeline (Mallett, 2016; Shollenberger, 2015). Considering the economic effects of dropouts alone, school suspensions in tenth grade generated social costs to the US totaling above $35 billion based on 67,000 associated school dropouts (Rumberger & Losen, 2016). In one tenth grade cohort of Californian students, a 6.5% drop in graduation rates due to suspensions alone (based on controlling for multiple other predictors of the state's 60% dropout rate) resulted in a statewide burden of $2.7 billion, of which $809 million came directly from taxpayers (Rumberger & Losen, 2017). Furthermore, students who are not suspended often experience academic decline and safety concerns when suspension rates rise in the schools they attend (Skiba & Rausch, 2006). When academics are linked to the discipline gap, the academic achievement gap between Black and White students relates to the same collection of concerns as well (Gregory, Skiba, & Noguera, 2010; Losen, Hodson, Keith, Morrison, & Belway, 2015).

Some argue that exclusionary discipline is necessary to ensure the achievement of behaving students, but this argument is flawed (Losen et al, 2015). Skiba and Williams

(2014) found that Indiana schools with lower suspension rates had higher achievement rates despite poverty and race. Another Texas study found that in schools with nearly identical demographics, increased use of exclusionary discipline did not correlate with better standardized test scores (Fabelo, et al, 2011). Perry and Morris (2014) tracked over 17,000 never-suspended students for three years and found that higher levels of exclusionary discipline in their schools associated with lower math and reading scores for those students. Later, the same team found just one suspension was linked to decreased standardized reading test scores of nearly one standard deviation below that of never suspended students, and that discrepancies in exclusionary discipline explained 20% and 17% of the differences in reading and math scores, respectively, between Black and White students (Perry & Morris, 2014). In a six-year effort to initiate restorative practices and decrease the use of suspensions in Denver, Colorado schools; suspension rates decreased as standardized test scores increased (Gonzalez, 2015). This demonstrated that exclusionary discipline did not ensure academic success for students who behaved and may even lead to academic decline instead.

Unfortunately, discipline gaps in exclusionary discipline and use of exclusionary discipline overall remain functionally unchanged—despite greater understanding of the risks of exclusionary discipline usage (Losen & Skiba, 2010; Losen, et al, 2015). From 2006 to 2010 both the Black/White and Hispanic/White racial discipline gaps increased with overall K-12 OSS rates rising from 15% to 16% of Black students, remaining at 7% of Hispanic students, and declining from 5% to 4% of White students. These rates remained the same for Black and Hispanic students in 2012, but rose again to 5% of White students. Although this is indicative of a lessened racial discipline gap, only an

increase in White student OSS led to this change rather than decreases in exclusionary discipline for any of the groups studied. When broken into elementary and secondary levels, secondary schools showed more promise than elementary schools. A slight reduction in secondary schools' OSS existed for all races with an even slighter gap closure seen by the Black/White discipline gap decreasing by 0.7 percentage points and the Hispanic/White discipline gap decreasing by 0.8 percentage points. In elementary schools, OSS rates rose for all races with the Black/White discipline gap decreasing by only 0.3 percentage points and the Hispanic/White discipline gap remaining the same despite a 0.4 percentage point increase in suspensions for White elementary students. Some areas saw even more drastic widenings such as Albany, New York where respectively the elementary and secondary school Black/White discipline gap rose 2.7 and 4.0 percentage points and the Latino/White discipline gap rose 4.8 and 2.0 percentage points (Losen, et al., 2015). The school discipline gap remains a problem nationally and grows as a problem when considered in key areas of the country.

In their 2014 study, Skiba and colleagues found that school-level variables contributed more to the overrepresentation of Black students in exclusionary discipline (particularly OSS versus in school suspension (ISS)) than behavioral- and student-level characteristics. Controlling for school-level variables made the relationship between race and exclusionary discipline insignificant. Although the percentage of Black students in the school was the strongest school level predictor of the odds of receiving an OSS versus an ISS, the principal's perspective on discipline was a major school-level predictor of the odds of exclusionary discipline (OSS/expulsion versus ISS). "In schools in which principals expressed attitudes more favorable toward school exclusion, students were

significantly more likely to receive OSS (odds ratio = 1.38) and expulsion (odds ratio = 2.32) relative to ISS" (p. 657). This suggests that the principal has more control over the likelihood of exclusionary discipline than student or behavioral characteristics and may share some of the responsibility for racial discipline gaps. Skiba later suggests: "…those wishing to have a positive effect on reducing or eliminating racial disparities in discipline would be well advised to seek interventions that focus on school policies and practices—principal leadership, achievement orientation, and the possible contributions of implicit bias—rather than on the characteristics of students or their behavior" (p. 664). To either confirm or deny the presence of racially motivated—but likely unintentional—biases in the severity of discipline received by students of Color, an effect of implicit bias requires formal establishment.

The Kirwan Institute, a leading research group for studies of race and ethnicity, defined implicit bias as, "The attitudes or stereotypes that affect our understanding, actions, and decisions in an unconscious manner. Activated involuntarily, without awareness or intentional control. Can be either positive or negative. Everyone is susceptible" (Straats, Capatosto, Wright, & Jackson, 2016, p. 14). Whereas explicit bias presents an outward, conscious attitude with intentional and voluntary expression such as traditional racism, implicit bias often remains unnoticed by those expressing biases. Many times, implicit bias contradicts explicit bias and people act in ways they would neither endorse nor condone with intention. This kind of disconnect, namely dissociation, results in decisions inadvertently affected by bias in either positive or negative manners depending on the circumstances and direction of the bias. These possibly flawed decisions present a great level of concern to proponents of the social justice movement

(Equal Justice Society, National Initiative for Building Community Trust and Justice, National Association for the Advancement of Colored People).

Consider a jury purposefully screened for racial biases by lawyers prior to a racially-loaded trial. In questioning the jurors, the lawyers will ask about opinions of Black or White individuals, experiences with such people, and outward biases. Potential jurors honestly responding to the questions may not express any racial biases but given a decision, unconscious associations about race could influence their responses. When deliberating a verdict, this jury might express no explicit bias but still make a decision heavily influenced by implicit bias. The jury does not expressly act biased and likely tries to eliminate bias from their verdict, but the verdict still contains implicit bias possibly resulting in a more severe verdict for the defendant. In such theoretical situations, implicit bias results in negative impacts. Fortunately, research is developing with strategies for counteracting the effects of implicit bias, but realizing the presence and consequences of implicit bias is critical before accountability can occur (Straats, Capatosto, Wright, & Jackson, 2016).

Definitions of Terms

All abbreviations are defined in Table 1 in addition to parenthetical definitions on first mention.

- **Race** refers to the skin color of a person as perceived by other individuals. Although race typically refers to a broader representation of a group of people that identify in some common manner, this study aims to understand how bias impacts individuals based on the color of their skin. As such, race does not refer

to one's ethnicity or culture for purposes of the present study and is further definate as phenotypic racial stereotypicality in Chapter 3.

- **Students of Color** refer to students who identify as Black or Hispanic for the purposes of the present study. Although other students experience racial bias based on the color of their skin, available student race data is limited to student self-identification. Only the categories of Black or Hispanic typically identify students with levels of skin pigmentation associated with bias in others.

- White Students refer to students who identify as Caucasian for purposes of the present study. Although many students may appear White without self-identification as such, other categorizations may include students who are not clearly either of Color or White (i.e. Indian, Arabic), students who experience different kinds of racial or ethnic biases (i.e. Arabic), or a variety of students of varying perceived races (i.e. Multiracial).

- **Implicit Bias** refers to attitudes or stereotypes that uncounsciously and unintentionally influence human behavior in a positive or negative manner that sometimes does not reflect one's explicit, or endorsed, beliefs and values (definition adapted from: Straats, Capatosto, Wright, & Jackson, 2016 and Greenwald & Krieger, 2006). In this study, implicit bias is presented in the context of racial biases between White students and students of Color.

- **The Implicit Associations Test (IAT)** is a measure of implicit bias created by Greenwald, McGhee, and Schwartz (1998). This test presents stimuli on a computer screen and asks participants to sort images and words as shown in Table

4. Error and response latency data are analyzed to create a score which represents the degree preference for the target stimuli.

- **Discipline** refers to the consequence(s) received by a student in response to a behavioral infraction, usually following an office disciplinary referral (ODR). Discipline is decided by a principal, assistant principal, or dean of students in PA K-12 public schools.

- **Discipline Severity** refers to the level of discipline in response to a behavioral infraction received by the student progressing from least severe to most severe in the following manner: warning, parent/student conference, detention, loss of privileges, weekend detention, ISS with education, ISS without education, OSS, alternative school placement, expulsion, arrest.

- **Exclusionary Discipline** is any disciplinary action which removes a student from the classroom or school building. This includes ISS, OSS, alternative school placement, expulsion, and arrest.

- **Suspension** refers to both ISS or OSS. In PA, OSS is limited to 10 school days per incident before it is considered an expulsion (22 Pa. Code § 12.6).

- A **Behavioral Infraction** refers to any student behavior which results in administrative discipline, usually by means of an ODR.

- **Subjective/Objective Infractions:** In the state of PA some disciplinary responses are mandated by law or school code of conduct in response to specific behavioral infractions, while other responses are at the discretion of the administrators (Safe Supportive Learning, 2016). These mandatory responses are considered objective infractions and by law include: possession of weapons; possession, use, or sale of

10

controlled substances; possession, use, or sale of alcohol or tobacco on school

property; violence; criminal activity or attempted criminal activity; and terroristic

threats. Subjective infractions include cheating, disruptive behavior, inappropriate

displays of affection, disobedience, property destruction, dress code violations,

profanity, missing detentions, and similar infractions. Infractions are designated

as subjective or objective in Table 2.

- **Infraction Level:** Behavior infractions are typically leveled based on severity by

  school districts in either three or four levels, where level 1 infractions disrupt the

  learning environment; level 2/3 infractions violate specific school rules; and level

  3/4 infractions put other students in immediate danger. As the level of the

  infractions increase, disciplinary actions typically become more defined by

  district policy or law.

- **Administrators** refer to the individual(s) in a school that make disciplinary

  decisions with regards to behavioral infractions. These individuals are usually

  school principals, assistant principals, or deans.

- **Socioeconomic Status (SES)** refers to the environment related to family income

  level experienced by students and is measured by eligibility for free or reduced

  lunch (FRL).

- **The School (Racial) Discipline Gap** refers to the discrepancy between the rate at

  which White students and students of Color experience exclusionary discipline.

  Losen (2015) defines multiple discipline gaps between students by race (Black-

  White or Hispanic-White) and SES. As such, this study focuses on students of

11

Color versus White students and controls for the effects of SES to investigate only the school discipline gap as it relates to race.

Study Purpose

The purpose of this study is to understand whether implicit bias influences the decision-making of PA K-12 administrators when making decisions about discipline in response to ODR for subjective infractions. Administrators' implicit bias should not affect objective infractions, those with disciplinary consequences guided by law or school policy, because no true decision-making should occur in this process. In discovering the predictive value of implicit bias on discipline severity by race, this study hopes to establish whether a need is present to train administrators on how to counteract the influence of implicit bias on their decisions—especially those that relate to discipline for students. Moreover, understanding and acting on this school-level social justice concern for discipline might extend into future judiciary concerns via the same mechanisms as the School to Prison Pipeline.

Significance of Study

The findings of this study will present a potential to further our understanding of the school discipline gap as it relates to social (in)justice. Implicit bias is a popular topic not only in research, but in modern events and politics as demonstrated by Former President Barack Obama's reference to implicit bias in his eulogy for Rev. Clementa Pinckney and others lost in the shooting at the Emmanuel African Methodist Episcopal Church: "Maybe we now realize the way racial bias can infect us even when we don't realize it, so that we're guarding against not just racial slurs, but we're also guarding against the subtle impulse to call Johnny back for a job interview but not Jamal" (Obama,

2015). With a nation beginning to recognize that bias can go unrecognized, showing a link between implicit bias and school discipline now may offer a platform for positive change instead of unproductive stagnation. If some of the school discipline gap's discrepancy is related to administrative implicit bias, then administrators could undergo professional development focused on reducing the effects of such bias. Possible methods include exposure to counter-narratives (Solórzano and Yosso, 2002) and/or counterexamples (Banaji, 2013; Dasgupta & Asgari, 2004) of typical racial stereotypes and mindfulness practices, both linked to decreases in implicit bias (Straats, Capatosto, Wright, & Jackson, 2016). Subsequent study of these interventions would evaluate efficacy as it relates to lowering school racial discipline gaps.

Findings from this study also show promise for developing more effective disciplinary systems in schools. School could implement systems such as Culturally Responsive Positive Behaviors Support which aim to:

(1) proactively reduce discipline problems by culturally and educationally supporting students in making desirable behavioral choices (Vincent, Randall, Cartledge, Tobin, & Swain-Bradly, 2011),

(2) develop disciplinary guidelines to reduce the influence of bias,

(3) lessen the cognitive load of subjective decisions to allow for more explicit bias mediation, and

(4) build teams of decision makers with varying levels of implicit bias to buffer more extreme bias.

Schools might consider implicit bias when developing school rules and policy with regards to discipline to reduce any cultural collision or collusion (see Beachum &

McCray, 2011) leading to inequitable codes of conduct. At a minimum, findings may persuade administrators or administrator-training programs to follow theory-based suggestions to include the IAT as an important element of professional development used to acquire a self-understanding of possible biases (Clark, & Zygmunt, 2014; Zygmunt, & Clark, 2016).

Definitions of Variables

- **Discipline Severity:** The severity of consequence(s) resulting from a behavioral infraction scored on an ordinal scare where: warnings and parent/student communications receive a one (1); weekday detentions, classroom removals, and loss of privileges receive a two (2); weekend detentions, alternative classroom placements, and fines receive a three (3), ISS receives a four (4), OSS receives a five (5), and alternative school placement, expulsion, and arrest receive a six (6).

- **Infraction Type:** The kind of infraction (objective or subjective) as defined earlier and demonstrated in Table 2.

- **Administrator Race:** The racial identity (Black or Hispanic, White, or Other) of the administrator making the discipline decision.

- **Administrator Experience:** The years of experience in the current position of the administrator making discipline decisions.

- **Administrator Implicit Bias:** The IAT score (see Chapter 3) of the administrator making the discipline decision.

- **Student Infraction Level:** The level of infraction on an ordinal scale where: behaviors disrupting learning for oneself or a class receive a one (1); behaviors disrupting learning for the whole school receive a two (2); behaviors which

14

verbally abuse students receive a three (3); behaviors which verbally abuse staff
receive a four (4); behaviors that physically harm students receive a five (5);
behaviors that physically harm staff receive a six (6); behaviors involving tobacco
or alcohol receive a seven (7); behaviors involving illegal drugs receive an eight
(8), and behaviors that endanger students (including weapon possession) receive a
nine (9). Infraction level is linked to PA Safe Schools codes in Table 2.

- **Student Grade:** The grade level of the student receiving discipline.

- **Student SES:** The FRL eligibility of the student receiving discipline.

- **Student Race:** The racial identity (Black or Hispanic, White, or Other) of the
  student receiving discipline.

- **Racial Discipline Discrepancy (RDD):** The discrepency percentage based on the
  *rati*o of students of Color receiving exclusionary discipline as decided by a given
  administrator to all students receiving exclusionary discipline as decided by the
  same administrator subtracted from the ratio of students of Color in the school
  where the administrator is employed to all students enrolled in the school where
  the administrator is employed. Per Reschly (1997) a group is over- or under-
  represented if exceeds representation in the relative population by $\pm 10\%$. As such,
  RDDs greater than 10% are considered over-represented and those less than -10%
  are considered under represented.

$$\text{RDD} = \left(\frac{\text{students of color recieving exclusionary discipline}}{\text{total students recieving exclusionary discipline}} - \frac{\text{students of color in school}}{\text{total students in school}} \right)$$
$$* \, 100$$

Research Questions

Discipline Severity

When considering student discipline separately for subjective, objective, and all decision types,

1. Are there differences in discipline severity between administrators? (RQ1)

2. Are differences in discipline severity by student race due to administrator implicit bias? (RQ2)

   a. …even when you take students' SES, grade, infraction severity and administrators' experience and race into consideration? (RQ3)

   b. Is this relationship the same for overall, subjective, and objective discipline?

Racial Discipline Discrepancy

1. Are discipline gaps based on race related to administrators' levels of implicit bias? (RQ4)

CHAPTER 2:

Literature Review

The purpose of this study was to understand whether implicit bias influenced school administrators' discipline severity decisions for subjective and/or objective infractions. Multiple studies established the existence of major inequities in ODRs and exclusionary discipline by ethnicity with students of Color receiving more OSS and expulsions than their White peers (Fenning & Rose, 2007; Noltemeyer & Mcloughlin, 2010a; Noltemeyer & Mcloughlin, 2010b; Mcloughlin & Noltemeyer, 2010; Skiba, Michael, Nardo, & Peterson, 2002; Townsend, 2000). These exclusionary discipline practices contribute to the School to Prison Pipeline phenomenon by increasing the likelihood of juvenile delinquency and subsequent failure (The Civil Rights/Advancement Project, 2000; Fabelo, et al., 2011; Skiba, Michael, Nardo, & Peterson, 2002).

Wu (1980) and Skiba et al. (2014) established links between administrator's beliefs and rates of exclusionary discipline. With the potential to develop a better understanding of factors involved in inequitable discipline by race, such relationships may inform policy, praxis, and professional development aimed to decrease racial inequities. As such, this literature review focuses on what is currently known about exclusionary discipline decisions made by administrators in an effort to understand the factors that influence such decisions. A thorough review of the research follows focusing on (1) detrimental effects of exclusionary discipline, (2) inequities in the rates of exclusionary discipline by race, (3) administrators' influence on exclusionary discipline

17

rates, (4) behavioral decisions particularly subject to bias, (5) overall implicit bias, and (6) implicit bias as it relates to schools.

Detrimental Effects of Exclusionary Discipline

As early as 1975, researchers discussed the detriments of exclusionary discipline: Suspensions: (1) take away educational time that may cause marginal, weak, or poorly motivated students to drop out permanently; (2) label children as "troublemakers" thereby making repeated behavior problems more likely; (3) deny children needed help, and (4) contribute to juvenile delinquency by putting unsupervised children and those with problems into the streets. (Children's Defense Fund, 1975, p. 62)

Nevertheless, exclusionary discipline practices continue and researchers continue to describe detrimental effects resulting from excessive exclusionary discipline use such as reduced instruction time (Losen, Sun, & Keith, 2017), academic decline (Morris & Perry, 2016; Perry & Morris, 2014; Rausch & Skiba, 2005; Skiba & Rausch, 2004); placement in the School to Prison Pipeline (Fenning & Rose, 2007; Mallett, 2016; Nance, 2016; Wald & Losen, 2003), high school dropout (Balfanz, Byrnes, & Fox, 2015; Costenbader & Markson, 1998; Marchbanks, Blake, Smith, Seibert, & Carmichael, 2014), social challenges (The Civil Rights/Advancement Project, 2000; Perry & Morris, 2014; Flanagain, 2007), grade retention (Marchbanks, Blake, Booth, Carmichael, Seibert, & Fabelo, 2015; Safer D. J., 1986), and even illegal drug use (Schwartz & Wirtz, 1990). Furthermore, exclusionary discipline practices appeared ineffective as shown by high incidents of repeat behaviors (Fenning & Rose, 2007; Safer, Heaton, & Parker, 1981) and

no decreases in rates of exclusionary discipline (Losen & Skiba, 2010; Tobin, Sugai, & Colvin, 1996). Still, school administrators use exclusionary discipline almost every day.

Reduced Instruction Time

Exclusionary discipline inherently removes students from the classroom or other learning environment, which reduces instruction time. On average, students in MA missed 16 days of instruction per 100 students enrolled. When divided by student race, Black students missed 34 days per 100 students enrolled while White students missed 10 days per 100 students enrolled (Losen, Sun, & Keith, 2017). Even when removing violent, drug-related, and criminal infractions; White students only averaged 6 days while Black students averaged 21 days of missed instruction due to discipline per every 100 students enrolled. Scott and Barnett (2004) estimated that suspended students in one urban elementary school missed 462 hours of instructional time in just one year. Considering that schools often suspend students for missing instructional time namely truancy or skipping classes, schools must consider instructional time important, but fail to demonstrate this when directing disciplinary measures.

Academic Decline

High levels of exclusionary discipline were associated with academic decline for both students receiving discipline and their incident-free counterparts (Perry & Morris, 2014; Rausch & Skiba, 2004, 2005). Three studies used state-wide datasets and correlation or ordinary least squares (OLS) regression to determine the relationship between academic success and exclusionary discipline. Rausch and Skiba (2004) used descriptive statistics to show that schools ranking in the bottom 25% of OSS versus the top 25% of OSS had 14.46% more students passing state-wide standardized tests. They

went on to determine that OSS rates explained 34.4%, 13.4%, and 39.2% of the variability in the percentage of students passing the state tests in all schools, elementary schools, and secondary schools, respectively. In their 2005 study, Rausch and Skiba took these results a step further by adding SES and demographic predictors in a step-wise regression model. They found that even when controlling for socio-demographic variables, higher rates of exclusionary discipline predicted lower rates of students passing the state standardized tests. Furthermore, after accounting for the poverty rate of a school, OSS was the next strongest predictor of achievement even when race and school level (elementary or secondary) were included as possible predictors.

Perry and Morris (2014) continued the investigation of the academic effects of exclusionary discipline using a different state's dataset with standardized test scores in reading and math separately. Their more thorough study accounted for longitudinal changes using a series of quadratic regression models and including variables indicative of SES, special education needs, rates and kinds of infractions, demographics, and school climate. With these models and predictors, the authors concluded that high-levels (as opposed to low or moderate levels) of exclusionary discipline negatively affected reading and math state assessment scores for non-suspended students. This relationship upheld despite controlling for level of school violence, socio-demographics, poverty, school-level variables, number of disciplinary infractions, and even ISS. Undoubtedly, a link exists between exclusionary discipline and academics.

Placement in the School to Prison Pipeline

Mallett (2016) defined the School to Prison Pipeline as, "a set of policies and practices in schools that make it more likely for students to face criminal involvement

with the juvenile courts than to attain a quality education" (p. 1). This connects exclusionary discipline to prison when school discipline leads to juvenile justice:

> When school discipline actions lead to juvenile court referrals, it may result in adjudication and probation supervision. If the pipeline is not disrupted and the young person does not do well while on probation or while supervised by the court, additional harm often ensues, including detention and/or incarceration placement. Youthful offenders who are held in detention centers…and those placed in longer-term juvenile jail facilities include many young people whose difficulties began in the schools; thus, this results in a cycle that becomes self-sustaining. (Mallett, 2016, p. 3)

Considering students of Color experience much higher rates of exclusionary discipline, the School to Prison Pipeline disproportionally impacts students who already suffer from school discipline discrepancies with initiation into the prison system.

Nance (2016) highlighted schools' role in School to Prison Pipeline entry citing CRCD statistics from 2011-12: "According to the 2011-12 CRD Collection, although African American students represented 16% of the total student population, they represented 27% of students that schools referred to law enforcement and 31% of students subject to a school-based arrest" (p. 1066). Just as schools over-represent students of Color in exclusionary discipline, they appear to over-represent these students in criminal referrals.

In creating a frame for the School to Prison Pipeline Research Conference, Wald and Losen (2003) described several studies working towards understanding how and why students entered the juvenile justice system. Some of these studies investigated possible

predictors of justice systems entry. In a report jointly issued by the American Bar Association and National Bar Association (2001), the authors found exclusionary discipline and being held back in middle school were the best predictor of future female adolescent arrests. Carmichael and colleagues (2005) found that holding all other predictors constant, students involved in at least one disciplinary infraction were 23.4 times more likely to be juvenilely incarcerated; additional infractions increased this by 1.5% and each day suspended from school by 0.1%. Texas Appleseed (2007) reported that students placed in Texas disciplinary alternative schools (an exclusionary discipline where students are placed in an alternative school for students with high rates of delinquency) had five times the dropout rate of mainstream schools and that 80% of Texas inmates were high school dropouts. Other studies link high school dropout with exclusionary discipline as well.

High School Dropout

Beyond the correlations found in the Texas Appleseed (2007) dropout study, Marchbanks, Blake, Smith, Seibert, and Carmichael (2014) reported links between high exclusionary discipline rates and high school dropouts. In a 2014 study, the authors found one instance of school discipline increased the student's likelihood of dropout by 24%. Because most students in the population (60%) received discipline, the authors' finding may be inflated. Information including the identities of the students committing infractions may be more explanatory of these dropout increases than the discipline itself. Nevertheless, the authors made an important point about the economic cost of these dropouts resulting from discipline: "the increase in dropout rate is associated with between $750 million and $1.35 billion in increased costs and lost wages of the lifetime

22

of each cohort" (Marchbanks, Blake, Smith, Seibert, & Carmichael, 2014, p. 20). Given

that these cohorts were only from the state of Texas, these costs become drastically

higher when extrapolated to nation-wide costs.

In 2016, Rumberger and Losen investigated such costs nationally using data from

Educational Longitudinal Study of 2002 and Florida state data. By controlling for other

determinants of dropout in a tenth-grade cohort from each dataset, the authors determined

that suspensions accounted for an additional 67,000 high school dropouts. Rumberger

and Losen (2016) used the same model of economic impact as the Marchbanks (2014)

study that:

> compares the economic outcomes of high school dropouts and high school
>
> graduates over their working adult lifetimes, from age 18 to 65, in four areas:
>
> earnings, crime, health, and welfare…expressed as the lifetime differences
>
> between dropouts and graduates in: incomes; taxes paid; government spending on
>
> health, crime, and welfare; tax distortions; and productivity gains. (Rumberger &
>
> Losen, 2016, p. 11)

The authors expected the 67,000 dropouts in their cohort to incur a lifetime cost of $11

billion in lost tax revenue and over $35 billion in social costs—or over $100 billion if

generalized to all tenth-grade US students.

Again in 2017, Rumberger in Losen repeated the study using all 2011-12 tenth

graders in the California Longitudinal Student Achievement Data System, which

included all students and schools in the state of CA. They found that students with

suspensions had a graduation rate of only 60% as compared to never-suspended students

with an 83% graduation rate. Furthermore, they found a single non-graduate incurred

economic losses of approximately $579,820 over his or her lifetime—extrapolating to $2.7 billion for all dropouts due to suspensions in this study. In perspective, this fiscal estimate both confirmed the national estimates of the previous study and the gravity of dropouts related to suspensions.

Social Challenges

Economic costs alone do not fully explain the detriments of the exclusionary discipline dropout rate increase. Costenbader and Markson (1998) described dropouts as the result of social detriments associated with exclusionary discipline. The authors found students who had never been suspended had less socioemotional impairment than those who has been suspended in school. Students externally suspended from school showed even more socioemotional impairment. Furthermore, students who had been externally suspended suffered from significantly more difficulties with rule compliance and reported significantly less interest in school. These compelling findings must be considered with caution, however, as they are based only on self-reported measures on a multiple-choice survey. Flanagain (2007) used a qualitative approach to understand these social detriments. In a sample of ten previously suspended students, upon returning to the classroom four students thought they were treated differently by the teacher, three students were never allowed to make up work missed during the suspension, two students had angry sentiments, and seven students were never offered anger management. These conditions are not conducive to a socio-emotionally healthy return to the classroom.

Perry and Morris (2014) examined how exclusionary discipline affected the entire school population including those students without disciplinary infractions. Using extent data collected for the Kentucky School Discipline Project, the authors compared

standardized test scores with the average exclusionary discipline rates to the kinds of infractions occurring in schools. They found that schools with high exclusionary discipline rates reflected a toxic school environment with increased infractions and violence overall and decreased standardized test scores for non-suspended students. "Excessive exclusionary discipline may produce social psychological outcomes that endure well after the punishment itself, and well beyond the individual who is punished, interacting with behavior to shape meanings, perceptions, and actions" (p. 17). Although this study mainly focused on math and reading test score changes based on exclusionary discipline rates and student- and school-level characteristics, the effects found in the other variables expressed meaningful insight to the social implications of exclusionary discipline.

Grade Retention

Yet another domain subject to the effects of exclusionary discipline, grade retention correlated with exclusionary discipline both before and after the incident resulting in a suspension (Safer, 1986). In this key 1986 study, Daniel Safer used the school data folders of 93 multi-suspended middle-school students as compared to 107 age and sex matched non-suspended middle school students who represented the average population (within one standard deviation) on a series of academic, attendance, mental/behavioral, and family factors. He found the number of times schools suspended a student positively correlated with retentions in both elementary and junior high school. Unfortunately, the students in this dataset were 91.5% male and no ethnographic data is included in the study, so results might not be generalizable to a larger sample. Nevertheless, in Flanagain's 2007 qualitative study, 50% of the five male and five female

previously suspended fifth graders were retained at least once (Flanagain, 2007).

Marchbanks and colleges (2014) found a student who was suspended even a single time

in ninth grade was 42.6% more likely to be retained at some point during junior or high

school. Furthermore, retention becomes an economic issue as the cost to educate a

student for one year is repeated an extra time. When this is added to delayed workforce

entry and loss of tax revenue from that student as a potential worker these costs skyrocket

(Marchbanks, et al., 2015).

Drug Use

In 1990, Schwartz and Wirtz conducted a study to help understand how to screen

for drug/alcohol abuse in adolescents through use of shortened version of the previously

validated Drug Alcohol Problem screening test. The screening measure was used to

evaluate 355 adolescents (ages 14-18) visiting a five-pediatrician office in Fairfax

County, VA (mostly White upper middle class). Although simply attempting to evaluate

the shortened screen for valid drug/alcohol abuse prediction, the authors also found that a

mere four items accounted for 70% of the predictive validity of the measure: (1) tobacco

product use, (2) accusation of having a drug/alcohol abuse problem, (3) school

suspension, and (4) riding with an intoxicated driver. The predictive validity related

solely to school suspension was not reported, but 53% of respondents with the highest

score for drug use (which was cautiously confirmed as accurate) reported school

suspensions. Despite a mostly White sample, these results add a health concern to the

already injurious practice of exclusionary discipline.

Exclusionary Discipline Inefficacy

Despite the many negative correlates associated with exclusionary discipline practices, many administrators continue to believe that exclusionary discipline effectively curbs problem behaviors and enhances school safety (Pudelski, 2014). In a 2000 content analysis of disciplinary codes, Fenning, Wilczynski, and Parraga reported disciplinary codes listed suspension as the most common response for all behavior infractions and 33% of the disciplinary codes reviewed listed suspension as the appropriate response for recurring tardiness (Fenning, Wilczynski, & Parraga, 2000). Rebecca Cohen (2013) suggested administrators see exclusionary discipline practices as a 'rationalized myth' or asocial purpose portrayed as technical purpose despite contrary empirical evidence (p. 6). In many ways, exclusionary discipline practices fit this definition in that they are repeatedly invalidated empirically but continue to be used in hopes of enhancing school safety.

Tobin and colleagues (1996) demonstrated that rather than deter problem behavior, suspensions increase those behaviors leading to further ODRs (Tobin, Sugai, & Colvin, 1996). Tracking several students through middle school, the researches saw increases in subsequent ODRs after a 6th grade start-of-year suspension when compared with students who were not suspended following a start-of-year ODR. Similarly, a study of students subjected to exclusionary discipline in one Florida district found, "The overall rate of recidivism was extremely high…only 31 pupils were one-time offenders; 75% of punished pupils committed one to five offenses during the year and 25% committed more than five offenses" (McFadden, Marsh II, Price, & Hwang, 1992, p. 144). Raffaele Mendez and Knoff (2003) replicated the study in the 12th largest school district in the US

and 2nd largest in FL, finding similarly high recidivism rates and concluded, "suspension alone often does not curtail inappropriate behavior" (p. 45).

Furthermore, school safety is not increased through exclusionary discipline, but instead is correlated with decline (Losen & Skiba, 2010; Skiba & Rausch, 2006). In fact, OSS and expulsion may contribute to increased community-wide crime: "Suspension and expulsion often provide troubled kids exactly what they do not need: an extended, unsupervised hiatus from school that increases their risk of engaging in substance abuse and violent crime" (Losen & Skiba, 2010, p. 11). Although a non-directional correlation, exclusionary discipline rates negatively correlated with safety ratings in schools despite control for community and school contexts (Steinberg, Allensworth, & Johnson, 2015). Considering the negative effects and inefficacies associated with exclusionary discipline practices, research needs to investigate why school administrators continue to widely use exclusionary discipline.

Inequities in Exclusionary Discipline

Documented since the 1970s, inequities by both race and gender exist in the utilization of exclusionary discipline (Anderson & Ritter, 2015; Children's Defense Fund, 1975; The Civil Rights/Advancement Project, 2000; Noltemeyer & Mcloughlin, 2010a, 2010b; Mcloughlin & Noltemeyer, 2010; Skiba, Michael, Nardo, & Peterson, 2002; Wu, Pink, Crain, & Moles, 1982). Students of Color, particularly African American students, and males, received significantly more exclusionary discipline than their White female counterparts in numerous studies. Despite investigations of other possible variables responsible for such inequities such as SES, more severe behavioral infractions, and statistical issues; these disciplinary disparities remain.

The Children's Development Fund published a pivotal report in 1975: *School Suspensions: Are They Helping?* This report evaluated data presented by the Office of Civil Rights to help determine if suspensions were effectively and equitably working towards better student outcomes. Unfortunately, much of what the report stated described negative outcomes and inequities associated with school suspensions. Regarding race, the authors acknowledged more suspensions occurred for White students at that time; however, the rates of suspension (how likely a student was to be suspended) were much higher for students of Color. Black students were three times as likely to be suspended as White students in elementary school and twice as likely in high school. Similarly, students with traditionally Hispanic surnames were suspended in elementary school at similar rates as White students, but nearly twice as much as White students in high school. Considering the recent desegregation of schools at this point, the results were as impactful then as they are now. Even gender proved a prominent area of disparity with male students receiving suspensions nearly twice as much as female students.

Despite reconfirming the higher rates of exclusionary discipline for male, Black, and low SES students; Wu and colleagues (1982) separated antisocial behavior and SES from race as contributors for disciplinary disparities (Wu, Pink, Crain, & Moles, 1982). In their study, the authors controlled for antisocial attitudes/behaviors and low SES both individually and together to try to lessen the variability explained solely by race. Unfortunately, neither variable removed the significance of race in explaining the variability in exclusionary discipline.

Our data clearly support the hypothesis that nonwhite and white students are not equally treated. The inequality in treatment exists even when factors such as

29

poverty, behavior and attitudes, academic performance, parental attention,

attending a centralized school, etc., are considered. To the extent this is true,

racial bias plays a role in suspension. (Wu, Pink, Crain, & Moles, 1982, p. 269)

This critical study opened the door to investigating racial differences from a civil rights

and social justice perspective by eliminating some of the rationale for bias-free racial

disparities.

In response to the zero-tolerance movement, The Civil Rights/Advancement

Project (2000) published a paper describing the detrimental effects and inefficacies of

exclusionary discipline related to this movement for all students, but particularly focused

on those of Color. When this study was published near the peak of the zero-tolerance

movement, African Americans students were 17% of those enrolled in public schools, but

32% of those receiving OSS. Furthermore, White students represented 63% of the

national student population, but only 50% OSS and 50% of expulsions. Twenty-five

percent of Black male students reported receiving exclusionary discipline at least once

during their four-year enrollment in high school.

Not only do Black students typically receive more discipline, but discipline is

more severe for Black students than White students committing the same infractions

(Nicholson-Crotty, Birchmeier, & Valentine, 2008). In a study of all Black and White

students between ages 10 and 17 in 53 Minnesota counties, the authors found that even

after controlling for a series of environmental factors Black students received more OSS

than White students for the same potential-OSS behaviors. The biggest gaps were for

violence and unspecified infractions, which typically had more subjective discipline

decisions. They further found larger exclusionary discipline gaps significantly associated

with higher juvenile referral rate differences between Black and White students—demonstrating another racially unbalanced entry into the School to Prison Pipeline.

Noltemeyer and McLoughlin (2010b; McLoughlin and Noltemeyer, 2010) looked at how school type and demographics related to disparities in discipline. Although they found higher levels of exclusionary discipline for males and students of Color in both studies, complementary findings of interest included the relationship of disadvantaged students, school type, and presence of African American teachers on the disparities. As the number of economically disadvantaged students in a district increased, the level of exclusionary discipline disparities by race decreased. As the number of African American students increased the number of suspensions increased, but as the number of African American teachers increased the number of suspensions decreased. Even when controlling for poverty, the rate of exclusionary discipline for Black students was two to three times higher for each kind of disciplinary infraction type as compared to White students. Furthermore, when controlling again for poverty, Black students received different levels of exclusionary discipline based on the school type (i.e. urban, suburban, rural, etc.) with urban, high-poverty school types demonstrating the highest level of disproportionality by race.

In another key study, Skiba, Michael, Nardo, and Peterson (2002) attempted to clarify possible causes of discipline disparities: statistical methodology, SES, and disproportionate rates of misbehavior. None of these appeared responsible for racial disparities in exclusionary discipline. In their sample of 11,001 middle school students with disciplinary infractions (out of a possible 50,000 student sample), the authors found a greater than 10% disparity between the referred and suspended groups and the total

31

sample for gender and race with males and Black student receiving more referrals and suspensions. Similarly, a greater than 10% disparity existed for expulsions by gender, race, and SES as determined by FRL eligibility. When investigating SES, the authors found that when controlling for SES, racial disparities in exclusionary discipline persisted; furthermore, the SES discipline disproportionality findings were less robust than those for race or gender. Finally, males were found to have higher incidences of severe behavior infractions than female—possibly relating to their increased referral rates. Conversely, students of Color had lower incidences of severe behavioral infractions as opposed to White students despite more frequent referral rates. Unfortunately, this study reinforced the possible race-based connection to exclusionary discipline disparities.

Skiba followed this study in 2014 with an investigation of what predicted an OSS versus an ISS. He found through use of hierarchical linear modeling that race and gender demonstrated the largest increase in odds for all student-level predictors (in that order) with Black students and males more likely to receive OSS than White students and females, respectively. Furthermore, when considering school-level predictors, Skiba stated:

> While neither behavioral nor other individual characteristics fully accounted for the contribution of race to OSS, school-level characteristics did reduce that relationship to non-significance. For racial disparities in suspension and expulsion, school-level characteristics appear to be more important predictors than behavioral or individual characteristics. (Skiba, et al., 2014, p. 658)

Apparently, more considerations about the roots of the disciplinary inequities deserve attention.

In 2014, Anyon et al. conducted a study similar to Skiba (2014) adding infraction and school level predictors to student level data to predict OSS odds rather than adding school- and principal-level data. Using district-provided data for all students enrolled in Denver, CO public schools, the authors conducted hierarchical analyses to compare contributors at various levels for risk and protective effects. They found Black and Multi-racial students were more likely to receive an OSS than White students for the same infractions despite controlling for various student and school demographics and interventions.

Anderson and Ritter (2015) attempted to investigate if racial disparities in exclusionary discipline continued to occur when accounting for a plethora of possible correlates such as infraction type, number of infractions, school level, school size, school demographics, and school region. They again could not account for racial disparities in discipline using other variables.

> … all else equal, being African American increases the relative odds of receiving OSS, expulsion, or referral to an ALE [alternate learning environment], and decreases the relative odds of receiving corporal punishment, no action, or "other" action relative to white students. In addition, being Hispanic increases the relative odds of receiving an "other (non-specified)" action and decreases the relative odds of receiving OSS, corporal punishment, or no action, relative to white students (Anderson & Ritter, 2015, p. 11)

Furthermore, they noted African Americans were more than six times as likely to be placed in an alternative learning environment than their white peers for the same kind of infraction. Although this decreased to 1.4 times more likely when controlling for school

level characteristics, these results indicated there was still very much a racial school discipline gap in the United States.

Administrator Influence on Exclusionary Discipline Rates

Administrators play a role in both the rates of exclusionary discipline and the different kinds of school disciplinary measures used. Studies described exclusionary discipline influence by administrators through disciplinary philosophies (Skiba, et al., 2014), experience and school climate (Christle, Nelson, & Jolivette, 2004), administrative centralization (Wu, 1980) or adherence to rule sets (Mukuria, 2002), and severity of response (Skiba R. J., et al., 2011). Nonetheless, researchers repeatedly find administrators have significant impact over rates of exclusionary discipline in schools.

Skiba et al. (2003) used the Disciplinary Practices Scale developed by Skiba and Edl (2004) to assess disciplinary philosophy differences effects on suspension rates. Schools with higher OSS rates had principals that tended to believe in zero-tolerance and suspension efficacy, but school principals with lower OSS rates tended to believe in suspension as a last resort, teaching appropriate behaviors, and adapting to student needs. Lower rates of African American suspensions correlated with principal beliefs of ISS viability, while higher rates of African American suspensions correlated with zero-tolerance beliefs. This measure was again used by Skiba et al. (2014) to investigate the likelihood of receiving OSS or expulsion versus an ISS. "In schools in which principals expressed attitudes more favorable toward school exclusion, students were significantly more likely to receive OSS (odds ratio = 1.38) and expulsion (odds ratio = 2.32) relative to ISS" (Skiba, et al., 2014, p. 657). Undoubtedly, principals' disciplinary philosophy beliefs related to exclusionary discipline in these studies; however, use of a limited

sample (Indiana only) and self-reported data potentially limit the generalizability of these results.

Christie, Nelson, and Jolivette (2004) took principals' beliefs a step further relating principal experience and school climate to exclusionary discipline rates. Principals with less experience (average of 4 years) came from schools with higher rates of exclusionary discipline than those with more experience (average of 11 years). In high-exclusionary discipline schools, only 27% of staff rated their school climate as "good" or better, whereas 100% of staff rated the climate as "good" or better in low-exclusionary discipline schools. Similarly, school personnel at high-exclusionary discipline schools cited poor family involvement, a need for resources, and a need to reduce suspensions as compared with low-exclusionary discipline school personnel citing "good" or better family involvement and little or no need for resources or lowering suspensions. Finally, researchers' observations of the school climate included instances of staff yelling at students in only high-exclusionary discipline schools. These findings suggested principals might lower suspensions by improving school climate and offer suggestions on how to begin based on staff and observational responses; however, differences might relate solely to principal experiences suggesting a need for superintendent involvement to enact positive change.

A key study by Wu (1980) measured administrative centrality based on school governance efficacy and how behavioral decisions were made: rules set by district, general policy, teacher policy, or unknown. Both poor school governance and more central rule administration correlated with increased exclusionary discipline rates despite control for student attitude/behavior, school level, and school location. Similarly,

Mukuria (2002) looked at the differences in exclusionary discipline rates based on school leadership in predominantly African American schools with particularly high or low exclusionary discipline rates. Principals in schools with lower exclusionary discipline rates:

> …followed the district suspension policy but did so with a contingency approach to discipline. The principals modified rules as they saw fit, depending on the circumstances. For example, they perceived the suspension policy as a flexible guideline but not a rigid document. (Mukuria, 2002, p. 441)

Principals in high-exclusionary discipline schools followed the district suspension policy like a legal code. One principal said, "The district policy is like a blueprint for me. It enables me to decide when I should suspend a student or not. Without it, it would be extremely difficult to make such a determination" (Mukuria, 2002, p. 442). The key elements to curtailing exclusionary discipline rates suggested by these studies are flexibility and good school governance.

When students commit minor infractions, disciplinary outcomes often become a choice for principals (Skiba, et al., 2011). When controlling for type/severity of infraction, the researchers found significant differences in disciplinary action by race.

> Both African American and Latino students are overrepresented in suspension/expulsion relative to White students at both K–6 and 6–9 levels. African American students are underrepresented in the use of detention at the K–6 level, and underrepresented in all administrative consequences except suspension/expulsion at the 6–9 level… The continuing significance of race/ethnicity…after controlling for type of behavior indicates that race/ethnicity

36

contributes to administrative decisions regarding discipline independent of type of infraction, beyond any prior disparity in classroom referral. (Skiba, *et al.*, 2011, p. 95).

Possibly most disturbing were the results found in elementary students: African Americans students were significantly more likely to receive an OSS for minor infractions and Hispanic students were more likely to receive all kinds of exclusionary discipline across every type of behavior infraction except disruption than White students. These results align with the Office of Civil Rights findings that Black preschoolers (18% of preschool enrollment) represent 48% of preschoolers receiving more than one OSS (OCR, 2014). Inequity in exclusionary discipline starts early, and principals appear to embody a power to contribute to racial discipline disproportionalities in a way which is not currently equitable.

Behavioral Decisions Particularly Subject to Bias

Not all behavioral decisions are equally subject to bias. In fact, teachers referred Black students to the office much more often for subjective offenses such as threat, excessive noise, and loitering rather than objective offenses such as fighting, carrying a weapon, or smoking (Bradshaw, Mitchell, O'Brennan, & Leaf, 2010; Skiba, Michael, Nardo, & Peterson, 2002). Black students more often received discipline for less-severe, discretionary offenses while White students received discipline for more-severe, non-discretionary offenses despite Black students receiving more ODRs (Fabelo et al., 2011; Kelly, 2010). Subjectivity in ODR decisions even translated into racial discipline gaps (Girvan, Gion, McIntosh, & Smolkowski, 2016). Using data from the SWIS online system, Girvan and colleagues revealed racial disparities in subjective ODRs accounted

for up to three times as much of the variability in exclusionary discipline discrepancies as compared to subjective ODRs. As controls were added, only subjective ODRs remained a significant predictor of exclusionary discipline discrepancies by race. In another study, the RDD of Black students was notable (greater than 10%) for only subjective discipline (Forsyth, Biggar, Forsyth, & Howat, 2014).

In their 2017 study focused on administrators' beliefs regarding exclusionary discipline and corporal punishment, Kennedy, Murphy, and Jordan contextualized this subjective-objective struggle using a qualitative design. In describing issues related to maintaining a strict discipline system one principal said:

> I guess I can't help but be a little subjective… it's different if you've built a relationship with the kid versus not having built one. They're easier to approach and almost easier to reprimand. I really don't believe it's human nature to be objective. (Kennedy, Murphy, & Jordan, 2017, p. 261)

Although focused on a positive subjective influence, this principal's admission of subjectivity presents a bias-filled approach to discipline. Another principal recounted:

> A referral the previous day for a student who was "disrespectful to teachers… throwing temper tantrums and acting out in class." He knew, however, that the student's mother had just been arrested the previous weekend and that a story about her had been on the front page of the local paper. The administrator struggled with understanding that the student's behavior likely stemmed from the student's "huge embarrassment" and "devastation" while also believing that he was obligated to punish the student for disrespect. Administrators struggled to balance honoring individual students' needs with maintaining consistency

38

because, as one noted, "you gotta have a heart" when working with students.

(Kennedy, Murphy, & Jordan, 2017, p. 262)

Again, despite great intentions, this subjective decision becomes wrought with threat of bias. The line between what the principal refers to as 'heart' and implicit bias is thin at best, and presents a challenge for equity—even with the best intentions.

Subjective decisions allow for bias to enter the decision-making process: "…social psychology research suggests that implicit racial biases are most likely to affect decision making when the decision involves an ambiguous situation and provides the biased decision maker some ground to justify the biased decision on nonracial grounds" (Simson, 2014, p. 545). A greater understanding of implicit bias may assist in furthering the scientific understanding of how and why minority students experience more exclusionary discipline—especially during subjective offense deliberations—and what can be done to lessen inequities.

Implicit Bias

Implicit bias refers to an unconscious bias that many display towards traditionally disadvantaged groups of people, and is operationally defined by scores on an IAT (Jolls & Sunstein, 2006). Despite attempts to disprove the existence of implicit bias, IAT scores link implicit bias to differences in:

- friendliness ratings (Fazio, Jackson, Dunton, & Williams, 1995),
- trustworthiness and trust decisions (Stanley, Sokol-Hessner, Banaji, & Phelps, 2011),
- racial profiling, law enforcement shooting behavior, and sentencing decisions (Banks, Eberhardt, & Ross, 2006),

- stereotypical memory errors (Levinson, 2007), and

- ratings of clinician care (Blair, et al., 2013).

Racial biases often develop as a result of a large collection of social and emotional influences including family, school, media, and community; as well as historical, experiential, educational, and political impacts. Considering that individuals often do not recognize the existence of implicit bias; the salience of these relationships cannot go unnoticed.

Implicit Versus Explicit Bias

Considering the pervasiveness of implicit bias, clarifying the delineation between implicit and explicit bias becomes critical. Overall bias is rooted in human beings' nature to categorize things and individuals to facilitate automaticity (Molenberghs, 2013). When in the context of in-group bias social categorization in the medial prefrontal cortex, action perception in the inferior parietal lobe, empathy in the anterior cingulate- and medial prefrontal- cortexes, and face perception in the fusiform area and amygdala (limbic system) build the neurological representative of bias. Together, neurological correlates for in-group bias lead to both implicit and explicit bias, but the neuroplasticity of the brain creates a mechanism for the malleability of both. Additionally, executive brain functions allow for control of behaviors related to bias. The neural basis of bias helps to clarify the natural and embedded system for bias in all people, but also exhibits the possibility of both conscious and subconscious bias in the inclusion of both typically conscious (prefrontal cortex, limbic system, anterior cingulate cortex) and subconscious (inferior parietal lobe, fusiform area) systems of the brain.

Greenwald and Krieger (2006) explained explicit biases, or beliefs, as:

A belief is explicit if it is consciously endorsed. An intention to act is conscious if the actor is aware of taking an action for a particular reason. Of course, actors may dissemble and deny they are taking an action for a particular reason, so conscious intentions based on explicit beliefs may be hard to verify. But a deceitful actor is nevertheless capable of asserting the belief or identifying the intention that provides the basis for action, even when unwilling to do so. (p. 946)

They contrast implicit bias as unconscious and unintentional with no control over, "social perception, impression formation and judgement" (p. 946). Based on this, explicit bias might reflect racism or sexism in the form of derogatory comments, hate crimes, or even just behaviors distancing oneself from others; meanwhile, implicit bias might reflect moving to a safe neighborhood (unknowingly because it is all White) or writing a review naming a store as rude or dirty (unknowingly based on the presence of more consumers of Color). In race, bias may connect to stereotypes such as Mexicans as lazy, Puerto Ricans as dirty, or African American as criminals. When explicit, individuals may simply refer to individuals in this way, but while implicit such stereotypes may inadvertently change behaviors such as walking down the side of the street where an African American is not walking.

The differences between implicit and explicit bias become less clear when considering the effects of either or both. In truth, the effects of implicit and explicit bias are often identical. If an individual considers an area where many individuals of Color reside unsafe, this could resonate from implicit bias, explicit bias, or both. The determinate lies in whether the individual consciously considered the area unsafe due to

the demographics of the neighborhood. This vagueness is a major confound in research because researchers often cannot measure explicit bias accurately.

When measuring explicit bias, individuals frequently fail to respond honestly likely due to the social inappropriateness associated with expressing such biases (Schuman, Steeh, Bobo, & Krysan, 1997). The most widely-used approximation for measuring explicit racial bias is the Modern Racism Scale (McConohay, 1986), which measures one aspect of explicit racial bias. Henry (2010) suggested modern racism most closely aligns with, "symbolic racism and racial resentment and is related to concepts such as subtle prejudice, racial ambivalence, and aversion racism" (p. 577). A more recent measure based on the Modern Racism Scale, the Symbolic Racism 2000 Scale (Henry, & Sears, 2002), offers a more-documented level of validity and reliability as well as a distinction from political conservatism often criticized as lacking in the Modern Racism Scale. Despite representing promising measures of one type of explicit racial beliefs, these tests do not measure basic explicit bias as a construct and fall short when assessing explicit bias in contrast to implicit bias (Dovidio, Kawakami, & Beach, 2003).

The Implicit Associations Test

Implicit bias is almost exclusively measured using the IAT developed by Anthony Greenwald and colleagues (Greenwald, McGhee, & Schwartz, 1998). In this test, contrasting visual stimuli such as male and female or Black and White faces are presented and a tester must categorize them on different sides of a computer screen. Next, opposite words such as "hello" and "goodbye" or a series of opposite-type words such as "unpleasant" words and "pleasant" words are presented and categorized on different sides of the screen. In the next round, the visual stimuli and words are categorized concurrently

with reaction time and errors recorded. Visual stimuli are presented again on reversed sides and the combined approach is subsequently repeated with words on the same side, but visual stimuli on the new side. Again, reaction time and errors are recorded. This sequence is depicted in Figure 1. The reaction time and errors from each combined session are analyzed to find a difficulty difference and this difference reflects implicit bias (see Chapter 3; Implicit Associations Test).

When taking the IAT for White/Black implicit bias, White Americans tended to score with pro-White attitudes even when controlling for facial familiarity (Dasgupta, McGhee, Greenwald, & Banaji, 2000). Furthermore, 50% Black Americans scored pro-White when an overall no-preference finding for Black Americans was further evaluated (Banaji, 2001). Even children show such pro-White preferences; when using a child-friendly IAT test, White Americans scored pro-White at ages six, ten, and as adults (Baron, & Banaji, 2006). Project Implicit, a research project at Harvard University, continues to collect data and reported findings of pro-White scores overall continue to exist.

> Results from this website consistently show that members of stigmatized groups (Black people, gay people, older people) tend to have more positive implicit attitudes toward their groups than do people who are not in the group, but that there is still a moderate preference for the more socially valued group. So gay people tend to show an implicit preference for straight people relative to gay people, but it is not as strong as the implicit preference shown by straight people. (Project Implicit, 2016, p. n.p.)

Consistent and clear pro-White implicit bias appears throughout the research, and unfortunately the effects of implicit bias do as well.

Implicit Bias in Socio-Economic and Workplace Interactions

Implicit bias is prevalent in workplace and social interactions (Jost, et al., 2009). Studies linked male/female implicit bias with differential hiring by potential employers who exhibited a preference confident and ambitious women less than men (Rudman & Glick, 2001) and increased dislike for women who succeeded in male-typical jobs (Heilman, Wallen, Fuchs, & Tamkins, 2004). Black/White implicit bias significantly predicted violent and racial slur use in college students (Rudman & Ashmore, 2007). In a study where a professional assistant of Color rated friendliness interactions with various individuals, IAT scores but not scores on the Modern Racism Scale were predictive of friendliness rankings (Fazio, Jackson, Dunton, & Williams, 1995). In another study, individuals rated the trustworthiness of individuals based on pictures differing only by race (Stanley, Sokol-Hessner, Banaji, & Phelps, 2011). Regression analyses revealed that IAT scores predicted trustworthiness ratings in that individuals tended to rate those belonging to the group they preferred as more trustworthy. Despite an 80% pro-white sample in this study, a bootstraps analysis confirmed the highly robust results.

These social effects extended to economic differences as well (Rudman & Ashmore, 2007; Stanley, Sokol-Hessner, Banaji, & Phelps, 2011). Rudman and Ashmore (2007) took their study a step further and included Jewish/Christian, Black/White, and Asian/White IATs. The participants were asked to make suggestions for budget appropriations prior to the IAT task and the relationship between economic trust and implicit bias revealed increased distrust for non-White, non-Christian ethnicities which

related to IAT scores robustly. Similarly, Stanley and colleagues (2011) had participants play a trust game where participants chose to give another "player" an amount of money ranging from $0 to $10, with the understanding that the partner would receive quadruple that amount. The other "player" was visually presented on a computer screen and was clearly either Black or White. The experimenter told the participants that the other "player" already chose to either give half of their money back to that participant or take all the money (make a mutual versus self-preserving economic decision). Regression analyses demonstrated a robust relationship between the amounts of money entrusted to the other players based on race and implicit bias as determined by the IAT. Unfortunately, implicit bias permeates into meaningful interactions in the social, workplace, and economic worlds.

Implicit Bias in the Legal System

Even before the courtroom, implicit bias affects in the US legal system. When police officers were asked, "Who looks criminal?" when presented with a series of Stanford University staff and students with no criminal record, they consistently cited those who were Black and those with more stereotypically Black features as the "criminals" (Banks, Eberhardt, & Ross, 2006). Furthermore, when both police and students were primed to think about violence, they tended to look at Black individuals more often than White individuals. But these biases go beyond simple racial profiling. Various studies have looked at shooting behavior in response to Black and White target individuals. One study used a video game to see where shoot or don't shoot errors were made (Correll, Judd, & Wittenbrink, 2002). Their results revealed:

45

…when the target was unarmed, participants mistakenly shot him more often if he

was African American than if he was White… When the target was armed,

however, participants mistakenly decided not to shoot more often if he was White

than if he was African American (Correll, Judd, & Wittenbrink, 2002, p. 1319)

A similar study presented either Black or White, police officers or criminals as the

shooting targets and participants were only supposed to shoot the armed criminals

(Greenwald, Oakes, & Hoffan, 2003). The results parallel those of the Correll study:

(a) subjects had greater difficulty distinguishing weapons from harmless objects

when the weapons were in the hands of simulated Blacks than Whites and (b)

subjects were response-biased in the sense of giving the weapon-appropriate

response more readily to Black than to White targets. (Greenwald, Oakes, &

Hoffan, 2003, p. 403)

Considering the prominence of race-based shooting and the "Black Lives Matter"

movement, these kinds of implicit bias-based errors cannot continue to go unnoticed.

Unfortunately, such errors extend into the courtroom.

Levinson (2007) presented judges and potential jurors with stories including

either a Black or White main character involved in a crime. Participants misremembered

the stories in ways where Black main characters were associated with more aggression

and violence than White main characters. Furthermore, this misremembering was not

associated with explicit bias as measured later in the study. Even sentencing data

reflected such bias with killers of White individuals receiving the death penalty much

more often than killers of Black individuals (US General Accounting Office, 1990). Even

when controlling for race and criminal history, the amount of Afrocentric facial features

predicted the length of sentencing for Florida criminals already incarcerated (Blair, Judd, & Chapleau, 2004). This indicated that despite the removal of race-based bias in sentencing, the more implicit stereotypicality remained of predictive relevance. If implicit bias can permeate our legal system, one should expect the infiltration of implicit bias on the school discipline system.

Implicit Bias in Healthcare

A recent meta-review of 42 studies (from 2003-2013) of implicit bias in healthcare confirmed a pro-white implicit bias in healthcare providers and negative correlations between level of implicit bias and quality of care in all the studies focusing on each of the topics (FitzGerald, & Hurst, 2017). To be included articles required: (1) an empirical design, a method of designating implicit rather than explicit biases, and (3) included physicians were not students. Despite quality control of the articles, the authors cautioned that publication bias (non-publication of non-significant results) potentially inflated findings. Although comparisons of implicit bias found in studied healthcare providers were like those exhibited in the general public, the sheer threat to wellbeing sprouting from such bias presents a highly concerning issue.

Additionally, implicit bias appeared to affect perceived levels of caring in the US healthcare system (Blair, et al., 2013). Patients surveyed about the level of caring provided by their clinicians reported less caring of clinicians whose implicit, but not necessarily explicit bias, was less favorable towards the patient's race/ethnicity. In fact, only implicit bias was related to the perceptions of care—possibly due to overall low levels of explicit bias. Implicit bias scores were variable and higher levels of implicit bias against the patient's race/ethnicity were related to lower ratings of clinician caring by the

patient. If these results replicate to principal-student or even teacher-student relations, one can only begin to fathom the possible effects.

Summary

Implicit bias is not only a real phenomenon (Jost, et al., 2009), but its effects permeate all areas of interactive society. From social interactions in the workplace, to economics, to legal and healthcare systems: implicit bias and its effects are supported in the research. Although implicit bias has not yet empirically made its way into the school administration literature, implicit bias is well documented in teachers (Warikoo, Sinclair, Fei, & Jacoby-Senghor, 2016). With implicit bias effects seen in judges, police officers, and doctors; the potential for implicit bias effects stemming from school administrators seems undeniable. Implicit bias appears to be a people problem, and with people being the core of the educational system in the US, one must expect implicit bias to work in similar ways within the school setting.

<div align="center">

Implicit Bias in Education

</div>

In their 2016 call for research and review of racial bias in education, Warikoo, Sinclair, Fei, & Jacoby-Senghor (2016) cited four reasons why implicit racial associations affect classrooms: (1) the pervasiveness of negative implicit associations toward people of Color, (2) the distinctness of explicit and implicit attitudes and mitigations focus only on explicit racism in schools, (3) correlations of implicit bias with, "problematic feelings and behaviors during interracial interactions" (p. 509), and (4) the typical conditions teachers work under which are vulnerable to the effects of implicit bias. All four of these contributors present legitimate concerns, which school implicit bias research has already confirmed.

Implicit Bias in K12 Teachers

Possibly the first study to demonstrate effects of implicit bias in teachers, van den Bergh and others (van den Berge, Denessen, Hornstra, Voeten, & Holland, 2010) looked at teacher bias in the Netherlands between Turkish/Moroccan students and Dutch students (a relationship that parallels the Black/White relationship in the US). The researchers measured both implicit and explicit bias using the IAT and Modern Racism Scale, respectively; and collected self-reported teacher expectation ratings, academic scores on national mathematics exams, and SES data. Using multi-level modeling, the team found a cross level interaction between implicit bias and race on teachers' expectations, but no effect of explicit bias on outcomes. A similar pattern was established for academic outcomes. This study established the ability of implicit bias to permeate into student success through teachers.

In a study of German pre-service teachers, Glock, Kneer, and Kovacs (2013) used the affective priming task to evaluate implicit bias towards immigrant and native students. This measure involved participants rating the affectivity of different words and then responding as either positive or negative to these words after presentation of images. Based on priming valence, participants should respond quicker to words with the same effect as the image presented prior, so if a picture presents a negative affect a subsequent negative word it should produce a quicker response than a subsequent positive word. The reaction times are compared to determine a positive or negative affect associated with each stimulus. In the study, the affective priming task revealed that pre-service teachers had a positive implicit bias for students "like" them and a neutral or negative implicit bias towards other students. In other words, the pre-service teachers preferred students who

looked similar to themselves. This finding highlights unconscious bias research considering racial match.

In a 2015 study, the Education Longitudinal Study of 2002 data were regressed to compare teacher expectations with student race (Gershenson, Holt, & Papageorge, 2015). The authors found that non-Black teachers had lower expectations of Black students with regards to future attainment (high school diploma or 4-year degree) than Black teachers, but similar expectations for students of other races. Relative to Black teachers, non-Black teachers were 12% less likely to expect a Black student to complete a 4-year degree. These effects were most prominent for Black males and from math teachers. Connecting this study to the Glock, Kneer, and Kovacs (2013) study, one may infer racial implicit bias accounted for such expectations.

Wright (2015) used data from the Early Childhood Longitudinal Study to apply racial congruence theory to discipline. With the extensiveness of this dataset, Wright compared the way a Black teacher saw a student to the way a White teacher saw the same student by controlling for average ratings across all teachers and the average rating for each teacher across all students—potentially allowing for causal statements. The findings of the study indicated that Black teachers found problem behaviors in Black males much less often than White teachers. Moreover, racial congruence did not matter for students of other races (i.e. White students with White teachers or Hispanic students with Hispanic teachers) or for females. The effects seen for Black males, however, were temporary and only existed for the time spent with that teacher (i.e. the year following a cultural match or mismatch, the new teacher-student match took over with no lingering effects of the previous teacher) suggesting a teacher bias issue rather than a behavior issue due to

behavior changes by year. The most notable finding of the study was cultural mismatch for Black males led to far more suspensions, and by simply doubling exposure to Black teachers for Black males the discipline gap might be halved. Even though this study does not explicitly measure bias, indications of bias and parallels to other studies create a place for this study in the implicit bias literature.

Cultural mismatch research lends itself to studies of the Cultural Synchrony Hypothesis which, "asserts that educators' negative evaluations of Black students are fueled by stereotypes of Black adults, who are depicted in the media as violent, threatening, hypersexualized, and in need of socialization. These negative evaluations have been shown to intensify when teachers do not share the racial/ethnic background of their students" (Blake, Smith, Marchbanks, Seibert, Wood, & Kim, 2016, p. 80). In a study of this latter statement, Blake and colleagues used Texas data to assess Black students' risk of discipline (defined by at least one instance of discipline during secondary school) based on the school level racial match or mismatch between teachers and students. They found that attending a school where teachers and students came from similar ethnic backgrounds benefited all students, but benefited students of Color to a far greater extent. Again, the study did not directly address implicit bias, but links to bias in the consideration of what about the match created a lower risk of discipline.

Glock and Karbach (2015) addressed links to implicit bias in teachers by measuring pre-service teachers' implicit bias using three different measures: the IAT, the affective priming task, and the affect misattribution procedure. The authors description of the affect misattribution procedure follows:

…this task does not rely on reaction times but rather on ratings of stimuli as

pleasant or unpleasant … This method assumes that the attitude object activates a

corresponding evaluation, which subsequently results in a judgment about a

Chinese pictograph that reflects this evaluation. If the attitude object is positively

evaluated, subsequently presented Chinese pictographs will be evaluated as more

pleasant than when the attitude object elicits a negative evaluation. (p. 56)

All three tests were presented in a random order to 57 German pre-service teachers who

scored as preferring or responding positively toward majority-race stimuli on all three

measures. Although this confirmed the presence of implicit bias in the sampled pre-

service teachers, findings of no significant correlations between the three test scores

presented a level of concern due to the intent of each assessment to measure the same

construct.

Employing a more empirical design, Okonofua and Eberhardt (2015) had 57

female K12 teachers imagine themselves teaching at a pictured middle school.

Subsequently, the researcher provided the teacher with the school disciplinary record of a

student named either Darnell or Deshawn (Black student) or Greg or Jake (White

student). All records were identical (except for counterbalancing the order of the

infractions) with two infractions: one for insubordination and one for class disturbance

with a short description of each. Next, the researcher asked the teacher to rate each of the

following on a one to seven scale: (a) the severity of the student's behavior, (b) the

hindrance of the behavior towards maintaining classroom order, (c) the level of irritation

experienced by the teacher, and (d) the severity of punishment appropriate. Afterward,

the research asked how likely the teacher would be to call the student a "troublemaker".

Finally, the teacher was asked a series of confirmation and control questions including the perceived race of the student and student and school demographics such as SES and school racial composition. Questions (a) through (c) were highly correlated and combined to represent a construct of "feeling troubled" and discipline was analyzed as recorded.

Results showed teachers were more likely to view multiple infractions as connected ("troublemaker"; see Townsend (2000) for more information on term use) for Black students as compared to White students, which mediated the relationship between student race and discipline severity. To confirm the findings, the researchers recruited and tested an additional 191 teachers. In this repetition, the researchers added two questions prior to the confirmatory/control questions: (e) the extent that the students' behaviors were indicative of a pattern and (f) whether the teacher could imagine suspending the student. The previous results were replicated. New analyses tested whether indication of a pattern mediated the relationship between student race and discipline severity and/or the relationship between student race and suspension; both mediating roles were found significant. This empirical study again did not remove the possibility of explicit bias as a contributor to these findings, but presented bias as a driver of perceptions that contributed to the race-discipline inequity.

Implicit Bias in Preschool Teachers

Gilliam and team (2016) demonstrated that bias is even present before students enter the K12 system (Gilliam, Maupin, Reyes, Accavitti, & Shic, 2016). In their two-part study, the team recruited and tested 132 early education teachers at a conference for teachers of early care and education professionals. The first experiment, presented teachers with a six-minute video of four preschoolers (one Black boy, one Black girl, one

White boy, and one White girl) playing at a table. Teachers were primed to look for problem behaviors in the students before they occurred and to press a key when a behavior was exhibited—no problem behaviors occurred in the video. Using eye-tracking software, the teacher's gaze was tracked and recorded for time spent looking at each student. After the video, teachers verbally identified the student they considered to need the most of their attention. Teachers spent significantly more time looking at the Black boy than other students and more time looking at the Black students than the White students overall. Similarly, teachers explicitly expressed the Black boy required most of their attention, followed by the White boy, then White girl, and finally the Black girl.

In the second experiment, teachers read a short vignette about a child's problem behaviors using female or male, Black or White names (Latoya, Emily, DeShawn, Jake, respectively). Half of the teachers were additionally provided with a student background as well. After reading the vignette (and background if appropriate), the teachers were asked to rate the following: (a) behavior severity, (b) degree of hopelessness for behavior improvement, (c) likelihood of recommending the child for exclusionary discipline, and if recommended for how many days. The participants rated the behaviors of White children as more severe, but there were no differences in consequence suggestions. Black teachers tended to recommend longer exclusionary discipline than White teachers despite student race. Teachers receiving background information rated the student as more hopeless. Without backgrounds, White teachers rated White children's behaviors as more severe than Black children's, but those receiving background information showed no difference in severity rating. Without backgrounds, Black teachers rated Black children's behavior as more severe, but those with background information rated White children's

54

behavior as more severe. Broken down, these findings revealed that the addition of background information increased severity ratings for students of a race different than the teacher. Not only do the findings of these two experiments confirm bias as early as preschool, but they offer a mechanism for implicit bias through looking for behaviors and interpretations of severity. Even when teachers explicitly reported spending more time looking for behaviors in White students than in the Black female students, actual behaviors reflected an implicit bias towards the Black girl. In the vignettes, racial mismatch led to increased ratings of severity when context was provided. Together these findings add a considerable concern to the school-based implicit bias literature.

Implicit Bias in Higher Education

Just as implicit bias extends into education prior to K12, the bias continues in post-secondary education. In 2015, Milkman, Akinola, and Chugh investigated college professors' responsiveness to students based on race. The team emailed 6,500 professors in 89 disciplines at 259 top US universities under the guise of a potential doctoral student inquiring about research opportunities on the path to graduate school. The team used identical emails signed with students' names randomly assigned to ethnicities including traditionally White, Black, Hispanic, Indian, and Chinese names to generate either a response or no response from professors. Results showed that White males received significantly more responses than all other groups across all discipline except for the fine arts. In the fine arts, White males received considerable less responses to the same degree as they received more in other disciplines. Discrimination gaps were highest in business, followed by education and human services. This study demonstrated a discrimination in higher education that might prove prohibitive for non-White male students' entry into

55

doctoral study. Furthermore, education was one of the highest affected areas which creates a potential for exclusion of highly educated, non-White or female individuals in education. This is especially concerning for potential principals and superintendents of Color who are already grossly under-represented.

In a review of higher education racial inequities in the United Kingdom, Alexander and Arday (2015) suggested that low rates of Black student acceptance and Black professional hiring at universities was related to implicit bias:

> British academia remains administratively, normatively, habitually and intellectually 'White', and Black academics and students suffer the most from the institutional racism and implicit biases that accompany this mono-culturalism. (p. 32)

The authors also discuss that implicit bias drives the creations of mono cultures: "As a result of unconscious or implicit bias, monocultures are created when people recruit in their own image. This is particularly true in senior positions" (p. 13). The issue related to mono cultures seemed to lie in the associated anti-Black sentiments:

> "…many Black academics feel themselves to be 'space invaders' operating in a predominantly White environment. Of course, being in a minority should not necessarily give rise to distressing experiences. However, the problem lies in the harmful racial/ gender stereotypes that are often held against isolated Black academics by mostly White senior colleagues and managers and expressed through a set of implicit biases. (p. 32)

Considering the almost hostile environment resulting from implicit bias, under-representation of people of Color in academia is not surprising.

In the US, even medical school admissions suffered from the effects of implicit

bias (Capers, Clinchot, McDougle, & Greenwald, 2016). The researchers in this study

gave the IAT to admissions committee members at Ohio State University's Medical

School. All participants demonstrated pro-White implicit bias with higher levels of bias

in men and faculty. After the test, 67% of the sample thought the IAT could reduce

admissions bias and 48% stated they consciously considered their results when

interviewing candidates for the next cycle. An additionally 21% of participants reported

that personal IAT result knowledge impacted their admission decisions for the following

cycle, which was the most diverse class admitted in history at the time of the study. Not

only did this demonstrate implicit bias in admissions, but it also highlighted the

malleability of the effects of implicit bias.

The Malleability of the Effects of Implicit Bias and Implicit Bias Itself

Despite the influence of implicit bias on education systems, this kind of bias

presents a solution more than a problem. Blair (2002) noted five areas of implicit bias

moderators; (1) self and social motives, (2) specific strategies, (3) focus of attention, (4)

stimulus cue configuration, and (5) individual category member characteristics (p. 244).

The first area, self and social motives, represents individuals' ability to change behavior

in order to preserve self-image or to comply with social norms. Specific strategies refer to

techniques such as counter-example exposure and suppression; while, focus of attention

indicates use of time and reflection to modify behaviors. Stimulus cue configuration

represents the heightening of bias based on stereotypical cues (i.e. a Chinese person using

chopsticks versus a Chinese person drinking a soda), and category member characteristics

meaning differences based on other associations related to an individual (i.e. a Black

friend versus a Black stranger). This key review, summarized nearly all the implicit bias malleability literature up to that date.

Newer studies provide even more implicit bias interventions, but still fit Blair's five areas. Measures of accountability can reduce the effects of implicit bias seen in teachers (Cate, Krolak-Schwerdt, & Glock, 2015). Teachers were tested with student vignettes like that of previously discussed studies at three timepoints (baseline, post priming, and 6-month follow-up). The teachers made tracking decisions and were held accountable at subsequent tests. The accuracy of the teachers' decisions increased with higher levels of accountability; specifically, after priming teachers about the importance of their tracking decisions differences in tracking decisions by race disappeared. Using the self/social motive of accountability completely eradicated the effects of implicit bias in this sample of teachers.

Counter-examples and mindfulness meditation represent modern implicit bias reductions strategies. Lai and several others (2014) examined 17 interventions entered in a contest of implicit bias interventions. They found interventions with counter-stereotypic exemplars were most effective at reducing bias in teachers. Dasgupta and Asgari (2004) similarly found that women surrounded by female leaders expressed less anti-female stereotypes and that the frequency of exposure mediated long-term effects of such exposure. Meanwhile, Lueke and Gibson (2015) found that listening to a mindfulness meditation as opposed to a controlled audio track decreased state implicit bias for both race and age. Stell and Farsides (2015) confirmed this effect with use of loving kindness meditation, which is similarly structured to the mindfulness meditation used in the Lueke and Gibson study.

In a study of an intervention program focused on implicit bias awareness and focused attention, Kahn, Goff, and Glaser (2016) demonstrated that a system including pre-intervention implicit bias measurement, a one day intervention, and post intervention measurement significantly reduced bias and related effects in police officers. The intervention phase consisted of implicit bias training, simulations of interactions vulnerable to bias where implicit bias mitigations strategies were used, training in the bias mitigation techniques, and then practice using those techniques.

In line with individual characteristic categorization, Mann and Ferguson (2015) conducted a series of six experiments evaluating whether implicit bias changes after new information prompted re-evaluation. In the first study, a narrative about a man breaking into a house was read followed by either the affect misappropriation procedure or IAT. Then, participants were told that the individual broke in to save children from a fire and the same test was administered. Regardless of test used, individuals significantly moved from negative to positive attitudes about the target after receiving the new information. The next study provided the same narrative and pre-test, but used a story about the man rescuing individuals on a subway instead. In this scenario, participants did not change their implicit bias related to the target individual. The third experiment replicated the first, but under conditions of a high, low, or no-cognitive load. They found a positive shift under each cognitive load level, but only full reversal in the low or no load groups. Experiments four and five replicated either the first or second experiment, with the addition of participant speed, deliberation, and participant-report of whether they re-evaluated the man after the new information. Findings based on stark differences in responses were interpreted to reflect completely different mechanisms for re-interpreting

versus re-evaluating, or elaborative thinking. The final study tested the longevity of reversed implicit bias finding that the reversal remained present three days after the original test. Together these experiments revealed that implicit bias was malleably long-term for a subject only when same-context new information was provided.

One study even brought multiple implicit bias remediation techniques together to develop a long-term implicit bias intervention (Devine, Forscher, Austin, & Cox, 2012). The researchers' intervention used training in stereotype replacement, counter exemplar imagining, individualizing, perspective taking, and contact over twelve weeks to produce long-term implicit bias reduction in participants. Reductions in implicit bias present at the fourth week of intervention were still present at week eight and possibly beyond. Furthermore, concern for people of Color at Week 2 moderated the relationship in that those exhibiting higher levels of concern developed greater decreases in pro-White implicit bias. This study holds much potential for implicit bias interventions for schools or other domains permeated by implicit bias.

Summary

When Carla Monroe (2005) coined the term, "School Discipline Gap," she described systematic differences between the rates of exclusionary discipline for Black and White students with the intentions of finding solutions. Over ten years later, we know more about how bad the problem is and little—if any—more about how to begin fixing it. Implicit bias is doorway into understanding the discipline gap and with research on how to lessen the effects of and presence of implicit bias increasing (Devine, Forscher, Austin, & Cox, 2012; Mann, & Ferguson, 2015; Kahn, Goff, & Glaser, 2016), it could also be the key to closing that gap. Skiba writes:

…those wishing to have a positive effect on reducing or eliminating racial

disparities in discipline would be well advised to seek interventions that focus on

school policies and practices—principal leadership, achievement orientation, and

the possible contributions of implicit bias—rather than on the characteristics of

students or their behaviors. (Skiba, et al., 2014, p. 664)

CHAPTER 3:

Methods

Design

The current study employed a non-experimental, cross-sectional design with data obtained from a target sample of the population during the first 100 days of the 2016-2017 school year. Non-experimental studies allow natural conditions to exist by using non-manipulated data and conditions (Cohen, Manion, & Morrison, 2011). Cross-sectional designs select a purposeful sample of the population with data collected for only one time interval (Cohen, Manion, & Morrison, 2011). Data collection occurred on a participatory basis through survey responses and extent dataset sharing. Participation invitations sent to schools and school districts included emails (Appendix A) to all school administrators inclusive of principals, assistant principals, assistant superintendents, and superintendents. If a school or district decided to participate, the researcher collected a signed informed consent (Appendix B) and site permission (Appendix C) form, subsequently shared with Lehigh University's institutional review board (IRB) as part of the research approval process.

After IRB approval, principals and assistant principals in participating schools and districts were invited to participate in a survey. Meanwhile, school data representatives shared or discussed extent datasets with the researcher to create a student-anonymous version of the data, inclusive of all necessary variables. After schools completed data entry for all discipline through the 100th day, data representatives shared these extent datasets with the researcher via email. Figure 2 depicts the data collection and IRB approval timeline.

<div align="center">**Population and Sample**</div>

The state of Pennsylvania contains 501 public school districts with a wide variety of rural, suburban, and urban communities. The 2011-12 OCR data show that the enrollment of Black students in PA schools (15.2%) approximates the national average (15.9%), but the enrollment of Hispanic or Latino students in PA (8.3%) is less than national average (23.6%). PA has similar incidences of exclusionary discipline for students of Color as national averages as seen in Table 3. As such, PA provides an ideal comparison demographic for studies of discipline and race with potential generalizability to the US school population.

Target Population

Of the 501 school districts in PA, 142 had student populations between 10% and 90% students of Color and at least ten reported incidents of discipline per the 2015-2016 PA public school enrollment data and Safe Schools Data available from the PA Department of Education. Due to the importance of student race with respect to the research questions, only school districts with between 10% and 90% students of Color were viable for inclusion in this study to allow for adequate variability in the sample. Similarly, the dependent variable related to instances of discipline, so the target population only included districts with at least ten reports of discipline. Not only did this provide a minimal level of data, but using districts with at least ten reported instances helped ensure valid data collection procedures existed in extent datasets for that district.

The target population included 140 school districts in the state of PA including representatives from all geographic locations in the state. Each district enrolled a minimum of 10% (Black: < 1%, Hispanic: < 1%) and a maximum of 90% (Black: 84%,

<div align="center">63</div>

Hispanic: 68%) students of Color, $\bar{x} = 29\%, SD = 20$ (Black: $\bar{x} = 16\%, SD = 17$;

Hispanic: $\bar{x} = 13\%, SD = 12$). The minimum number of disciplinary offenders for the

2015-2016 school year was ten and the maximum was 5882, $median = 66, \bar{x} =$

$567, SD = 185$. Per the 2011-12 OCR exclusionary discipline data, students of Color

were overrepresented in exclusionary discipline in 81% ($n = 117$) of the districts in the

target population and underrepresented in 1% ($n = 2$) of districts. White students were

overrepresented in exclusionary discipline in 1% ($n = 2$) of districts in the target

population and underrepresented in 78% ($n = 112$) of districts.

Actual Sample

Sixty-one schools nested in seven districts (of which four districts had only one

participating school) agreed to participate in the study. Of the 124 administrators in the

participating schools/districts, 41 administrators representing 27 schools completed the

survey. The final sample included 22 schools nested in seven districts. Although an

additional 39 schools agreed to participate in the study, 15 schools had no disciplinary

data to report and 34 schools had no participating administrators. Demographic data for

participating schools and districts are presented in Tables 6 and 8.

Response Rate

Low response rates at both the district and administrator levels were present with

rates of 5% and 33% respectively. Implicit bias and administrator-associated school

discipline data are both highly sensitive due to the potential for legal ramifications and

judgment. Furthermore, many districts noted concerns related to time and effort required

for full participation in the study. Finally, despite blinded student data and administrative

data linked only to an ID defined in a separate file, some districts raised concerns over

student anonymity and administrator confidentiality in their decisions not to participate. As such, these low response rates were not surprising.

Data Collection

The researcher emailed all principals, assistant principals, superintendents, and assistant superintendents from the 140 potential school districts to recruit participants. The invitation email (see Appendix A) detailed the requirements for study participation, benefits and risks of participation, and the purpose and goals of the study. Additionally, a pro-bono discipline equity audit (see Skrla, Scheurich, Garcia, & Nolly, 2004) using the district-provided data and an online bias training tool functioned as compensation for participation regardless of administrators' participation decisions. A discipline equity audit uses school discipline data including discipline severity/level, student race, special need status (presence of IEP or Section 504, socio-economic status (FRL eligibility), and other school-requested variables to determine if discipline is equitably administered to students. The equity audit compares the proportion of each subset of students experiencing discipline to the proportion of each subset of students in the school/district population to determine whether inequities exist. The discipline equity audits shared in the present study followed the template showing in Appendix B. The researcher sent the invitation email twice with a period of a 54 days (30 in-school days) between the two emails.

Student Behavioral and Demographic Data

Student behavioral data is regularly collected in PA public schools by state mandate for the PA Safe Schools database. Per 24 P.S. § 13-1301-A, school districts must report certain disciplinary incidents inclusive of: the age and grade of the student, the

address of the school, the circumstances surrounding the incident, the student's race, whether the student has an IEP, the disciplinary sanction imposed by the school, and any criminal charges. As this information is already collected by schools, this study hoped to gain access to the extent datasets.

The student behavioral data desired for purposes of this study included: infraction type/level; students' grade, race, FRL eligibility, and the disciplinary action with deciding administrator. Infraction type was binarily categorized as objective or subjective and infraction level rated on an ordinal scale from one to nine per Table 2. Student grade ranged from kindergarten (grade 0) to grade 12 and was recorded on an ordinal scale. FRL eligibility was categorized as eligible for free or reduced lunch (1) or not eligible (0). Discipline severity was scored on a 1 to 6 ordinal scale with warnings and conferences receiving a one (1); weekday detentions and loss of privileges receiving a two (2); weekend detentions, service, and fines receiving a three (3), ISS receiving a four (4), OSS receiving a five (5), and alternative school placement, expulsion, and arrest receiving a six (6). For discipline actions reported as subject to hearing, the research assigned a severity score of six due to the potential for high-level severity. Due to the intention of this data for use by schools and intended accessibility for any researcher, all ordinal data was treated as continuous for study analyses to remove the barrier created by the complexity of findings associated with logistical regression odds with more than 3 possible outcomes.

Due to suggestions from school districts, analyses used a three-level coding of race rather than binary coding (White and of Color) to account for the wider variety in skin color seen in Hispanic students as compared to African American students. Such

66

coding resulted in a measure approximating phenotypic racial stereotypicality (PRS) rather than strictly student-identified race (Kahn, & Davies, 2011). In PRS, race is considered with within group variability linked to higher and lower levels of experienced bias based on skin tone. Using a three-level continuum of race characterized by PRS, provided an ordinal scale of the likelihood of a student to experience bias based on the color of his or her skin for those students who typically had skin-colors associated with their racial identification. Race categorized by PRS is in line with the recommendations of Maddox and Gray (2002), Maddox (2004), and Relethford (2009) to focus on perceptions of skin color by phenotype rather than simply ethnicity. Furthermore, PRS-based measurements of race align to a more automatic bias paradigm based on amygdala responses based on skin tone (Ronquillo, et al., 2002). As such, student race was recorded on an ordinal scale as Black (2), Hispanic (1) or White (0), where Black/African American students were coded as Black, Hispanic/Latino students were coded as Hispanic, Caucasian students were coded as White, and all other student races were dropped from data analyses. As this study focused on bias as a function of perceived race, only races that typically associated with a PRS were viable for inclusion (e.g. Asian students might appear light skinned as a typical Chinese individual or dark skinned as a typical Indian individual). Similarly excluded, infraction data for students noted as multiracial or of two or more races led to unclear racial categorizations based on PRS.

For all administrators with at least 10 instances of exclusionary discipline, the researcher calculated the RDD with the following formula:

$$RDD = \left(\frac{students'\,of\,Color\,exclusionary\,discipline}{total\,students'\,exclusionary\,discipline} - \frac{students\,of\,Color\,in\,school}{total\,students\,in\,school} \right) * 100$$

Per Reschly (1997), a discrepancy exists when a subgroup's sample representation exceeds +10% of the total population. As such, RDDs exceeding +10 (any RDD outside of the -10 to +10 range) were considered discrepant with those greater than 10 indicating over-representation and those less than -10 indicating under-representation.

Administrator Demographics

As part of an online survey administrators were asked to share their race and years of experience in their current position just prior to initiation the IAT. Because administrator race represented only a control variable, binarily categorization as White or of Color remained appropriate. This categorization allowed for control related to whether the administrator likely experienced bias like students, involvement of race match/mismatch (Wright, 2015; Glock, Kneer, & Kovacs, 13), and as a control for the effects of cultural collision and collusion (see Beachum & McCray, 2011). Years of experience recorded continuously by the number of years in the current role accepted integer responses only. The survey also collected an administrator passcode purposed to link the administrator with his or her related disciplinary data. The researcher provided administrator passcodes and corresponding administrators' email addresses to schools to enter the discipline data prior to sending if possible in an effort to enhance confidentiality of the data.

Implicit Associations Test

The IAT provides a valid and reliable measure of implicit bias (Greenwald, McGhee, & Schwartz, 1998) as shown in a variety of assessments. Measures of reliability for the IAT demonstrated internal consistency with split-half correlations and alphas between .70 and .90, test-retest reliability with a median of $r = .56$, and an inability of

participants to fake results when directed to do so (Nosek, Greenwald, & Banaji, 2007).

Construct validity of the test is often weak as are nearly all measures of implicit

constructs (Nosek, Greenwald, & Banaji, 2007). In a study of seven measures of implicit

constructs including the IAT, Bosson, Swann, and Pennebaker (2000) found only weak

relations ($r = -.14$ to .23) between the measures. In discussing the construct validity of

the IAT with regards to the IAT measuring a different construct than measures of explicit

associations, Nosek, Greenwald, & Banaji (2007) write: "the best-fitting models

represented the IAT and self-report as related but distinct constructs, rather than as a

single attitude construct, even after accounting for common method variance in both

measures" (p. 278). The inter-item reliability of the measure is lower than is generally

acceptable for construct measures with a Cronbach's alpha of .69, but stability and

convergent validity overcome this problem (Cunningham, Preacher, & Banaji, 2001).

Cunningham wrote: "In two confirmatory factor analyses, each of the implicit measures

substantially and reliably correlated with the others, demonstrating convergent validity

for implicit attitude measures" (Cunningham, Preacher, & Banaji, 2001, p. 170).

Greenwald, Poehlman, Uhlmann, and Banaji (2009) performed a thorough meta-analysis

of studies using the IAT and found that those studies looking at social group

discriminations had explicit and IAT scores that predicted behavior with the IAT acting

as a better predictor (IAT: $r = .25$, Explicit: $r = .13$). Similarly, Greenwald and colleagues

(2009) found the average predictive validity of the IAT was $r = .27$ and the predictive

validity for the Black-White IAT was $r = .24$. The test manages such validity and

reliability using reaction time rather than typical self-report to measure the construct. The

measure controls for the effects of primacy with random ordering and the effects of

redundancy with reaction time matching and embedded outlier deletion. Additionally, the creators of the IAT re-evaluate the test regularly to ensure the highest potential for reliable and valid results.

Despite the reliability and validity of the IAT, the test is not without flaws. In their 2015, Blanton and Jaccard listed ten challenges to the scoring, use, and interpretation of the IAT. They first asserted the importance to assess explicit bias before measuring implicit bias. Although an important consideration, modern measures of racial explicit bias are limited to only symbolic racism (Henry, & Sears, 2002), which fails to consider all aspects of explicit bias as compared to implicit bias. Next, the authors considered the low convergent validity (relationship between various tests of implicit bias) suggesting that even tests showing the highest levels of convergent validity ($r =$ 0.45) are below the generally accepted minimum of 0.70 (Blanton, & Jaccard, 2015). The authors went on to express scaling concerns by comparing the IAT to a temperature scale where the interval between degrees was unknown. They cited a lack of consideration in the scoring algorithm for variable error and random error and suggested a lack of context for the test was of concern. Blanton and Jaccard noted that the IAT forced a composite score where preference and non-preference were interpreted of one score rather than providing a measure accounting for a dislike-driven preference where a strong dislike of one race might indicate a preference for the other presented race even if that race was also disliked. They went on to extend this concern into use of the IAT measure as a dependent variable because it might inaccurately measure implicit bias due to the single score. The authors mentioned a lack of control for participants' general processing speed, but acknowledged accountability for this with the modern scoring algorithm with a caveat

related to new concerns introduced with this algorithm. Oswald (2013) detailed this concern suggesting that using the in-group standard error in the equation caused measurement noise to increase the test accuracy rather than decrease as would be logical. Blanton and Jaccard (2006, 2015) and Blanton, Jaccard, and Burrows (2014) asserted that the IAT uses an arbitrary set of cutoffs to decide whether bias is small, medium, or larger; however, Greenwald agreed with this and noted that many tests of psychological constructs have arbitrary cut-offs (Greenwald, Nosek, & Sriram, 2006). In the same article, Greenwald and colleagues responded to nearly all of Blanton and Jaccard's ten challenges with an overall message that the test does not aim to diagnose, but only to assess—similar to tests of blood pressure. This analogy is again used in correspondence for a recent VOX.com article where Greenwald explained better aggregate than individual results for the IAT (Lopez, 2017). The most often cited complaint of the IAT remains very low effect sizes in results (Blanton, Jaccard, & Burrows, 2014; Oswald, Burrows, Blanton, Jaccard, & Tetlock, 2013; Bartlett, 2017), but Banaji and Greenwald (2015) write, "statistically small effects can have societally larger effect" p. 553). Despite notable concerns, the IAT using the modern scoring algorithm remains the most widely used and likely most accurate measure of implicit bias.

The modern version of the IAT is organized into seven blocks as displayed in Figure 3 (Greenwald, Nosek, & Banaji, 2003). When scoring the IAT only trials from blocks three, four, six, and seven are used. First data is screened for accuracy by eliminating any respondents where more than 10% of their responses include latencies over 300 ms. Next, trials with latencies under 400 ms or over 10,000 ms are removed from the data. Scoring begins by computing the mean of the correct-response latencies

71

for each block and computing pooled SDs for trials in blocks three/four and six/seven. Error-response latencies are replaced with the block mean plus 600 ms. These new values are used to average the block latencies once again, then the average of block three is subtracted from block six and the average of block four is subtracted from block seven (B6 – B3 and B7 – B4). These two differences are divided by their respective pooled SDs and the resulting quotients are averaged. This value is the IAT score and is interpreted in the same way as a Cohen's d value. When the later blocks include students of Color and positive pairings as shown in Figure 3, positive IAT scores indicate preferences towards White students and negative scores indicate preferences towards students of Color.

Presentation of the IAT occurred electronically on the participant's computer and at his or her leisure. During the test the images and words presented in Table 5 were used to designate White students and students of Color and the positive and negative terms. Participants pressed the "E" key on their keyboards to assign stimuli to the target on the left portion of the computer screen and pressed the "I" key to assign stimuli to the target on the right portion of the computer screen. The IAT and demographics data were programmed, presented, and recorded using PsyToolkit (Stoet, 2010, in press). PsyToolkit recorded the IAT response accuracy and latency and the researcher analyzed the collected data in MS Excel per the modern IAT scoring guidelines (Greenwald, Nosek, & Banaji, 2003) with the exception that responses less than 400 ms and greater than 10,000 ms were not accepted by the program. Instead, responses were locked until 400 ms and after 10,000 ms the program marked an incorrect response and progressed to the next stimulus. Despite this built-in response latency control, no participants attempted to respond before 400 ms or failed to respond prior to 10,000 ms.

District/School Demographics

Districts and school demographics were obtained using data available on the PDE's website for enrollment and personnel reflecting data as of October 1, 2016. Behavioral data were obtained from the PA Safe Schools database reflecting data through December 31, 2016. In addition to acting as paradata when testing for nonresponse bias, these data were used in the calculation of RDDs. Determination of the RDD subtracted the administrator specific percentage of exclusionary discipline experienced by students of Color from the percentage of students of Color in the administrator's school.

Data Analysis

For each dataset (i.e. Administrator, Subjective Discipline, Objective Discipline, All Discipline) descriptive and correlational analyses were run in SPSS. Data were analyzed using a combination of HLM and regression. Both HLM and OLS linear regression use the General Linear Model ($\bar{y} = X\bar{\beta} + \bar{\epsilon}$) to statistically evaluate relationships between variables (Neter, Kutner, Nachtsheim, & Wasserman, 1996). As such, these statistical tests were best suited for the research questions. Due to the presence of nested data (i.e. student discipline nested in administrators), HLM was the more appropriate statistical method as compared to OLS linear regression for discipline severity predictions (Raudenbush, & Byrk, 2002). Investigations of RDD only occurred at the administrator level and did not involve nesting, which allowed for the use of simple linear regression.

The researcher screened raw data for illogical values and entered valid data into IBM SPSS Statistic V24.0 with administrator data in one file and student discipline data in a second file. Due to multiple discipline instances per student and no reliable method

73

of determining the instance order for all students, data were randomly selected for inclusion of one administer-student case for each unique set. Administrator IDs and Student IDs were combined in the student data to create a unique ID for each student-administrator data relationship in addition to an assignment of a random ID number between 1 and 500,000 for each case. After sorting by the random IDs, the researcher used the Identify Duplicate Cases tool (last case primary) to create a unique/primary case filter variable for each student-administrator ID and subsequently saved all selected unique/primary cases into a new file using the Select Cases tool. After adding a preselection step to delete either subjective or objective infractions from the data to create subjective- and objective-only student discipline datasets including the most cases possible, the random selection process repeated.

Regression analyses were performed in SPSS using the enter method. HLM analyses were conducted in in HLM 7.01 (Raudenbush, Byrk, & Congdon, 2013) to determine the appropriateness of multilevel modeling and conduct analyses if deemed appropriate. All HLM analyses were run using the Restricted Maximum Likelihood Estimation method and the HLM2 design. Full Maximum Likelihood estimation was not used as it is only appropriate for comparisons of nested models (Raudenbush, & Byrk, 2002). Descriptive statistics obtained using SPSS were confirmed in HLM 7.01. When conducting HLM analyses, the researcher ran models in the following order using student discipline severity as the dependent variable:

1. Fully unconditional model (FUM, empty/null model)
2. Hypothesized model: FUM with the added predictors of student race/PRS and administrator implicit bias only.

3. Controlled Model: Model from (2) with addition of control variables (fixed effects) in the following order: student infraction level, student FRL eligibility, student grade, administrator race, administrator experience. If a variable was significant at the 0.05 level it was retained as the next variable was added to the model. If a variable was not significant it was dropped before the subsequent variable was added to the model.

4. After all control variables were evaluated, the final model was re-run to create residual files for assumption analyses.

After all models were run in HLM 7.01, pseudo R^2 values were calculated on significant findings to determine effect sizes using the Snijders and Bosker (1994, 2012) method (1 -

$\frac{\sigma^2_{Model}+\tau_{Model}}{\sigma^2_{FUM}+\tau_{FUM}}$ or 1 - $\frac{(\frac{\sigma^2_{Model}}{HM})+\tau_{Model}}{(\frac{\sigma^2_{FUM}}{HM})+\tau_{FUM}}$, HM = Harmonic Mean). This method defines,

"measures of modeled (or explained) variation [by] the principle of proportional reduction of prediction error" (Snijders, & Bosker, 1994, p. 351) and provides a more stringent calculation with regards to unbalanced designs. This calculation occasionally provides invalid results due to negative findings. When this occurred, pseudo R^2 calculations using the formulae suggested by Kreft & de Leeuw (1998) and Singer (1998)

were attempted ($\frac{\sigma^2_{FUM}-\sigma^2_{Model}}{\sigma^2_{FUM}}$ or $\frac{\tau_{FUM}-\tau_{Model}}{\tau_{FUM}}$). Although the Kreft & de Leeuw

(1998)/Singer (1998) method provided a less complicated calculation of proportioned variance, the formula did little to account for unbalanced design seen in this study making the Snijders & Bosker (1994, 2012) model preferable. Due to the manual nature

of these computations using a calculator and MS Excel were employed separately to minimize the risk of calculation errors.

Missing Data

Missing data were only present for non-participants (i.e., administrators in participating schools that chose not to participate) and for non-applicable data (i.e., multi-racial or other race/ethnicity students). Missing data at the student level were deleted listwise when running analyses, but at the administrator level entries with missing data were deleted listwise when creating statistical files for HLM 7.01 due to software requirements. For RDD analyses using single level OLS regression, missing data were also deleted listwise when no RDD was available.

Research Questions with Models

Each research question addressed with a different prospective HLM or regression model with fixed-only effects functioning as controls and fixed and random effects functioning as the independent variables of interest. Fixed effects in HLM are similar to constants, or intercepts, in OLS linear regression and describe the predicted intercept or mean of a DV based on a particular IV. Random effects approximate regression coefficients, or slopes, in OLS linear regression and describe the predicted slope or change in the DV for every unit increase in the IV. By fixing effects in HLM, a researcher chooses to not allow the IV to vary at level 2; allowing for random effects assumes the effects related to that IV will differ across level 2 units.

Comparing the findings for subjective, objective, and all instances of student discipline separately:

1. *Do differences exist in discipline severity between administrators? (Fully unconditional model)*

   Level-1 Model: $DISSEV_{ij} = \beta_{0j} + r_{ij}$

   Level-2 Model: $\beta_{0j} = \gamma_{00} + u_{0j}$

   Mixed Model: $DISSEV_{ij} = \gamma_{00} + u_{0j} + r_{ij}$

2. *Does administrator implicit bias adjust the relationship between student race/PRS and discipline severity? (Hypothesized Model)*

   Level-1 Model: $DISSEV_{ij} = \beta_{0j} + \beta_{1j}*(RACE_{ij}) + r_{ij}$

   Level-2 Model: $\beta_{0j} = \gamma_{00} + u_{0j}$

   $\qquad \beta_{1j} = \gamma_{10} + \gamma_{11}*(IAT_j) + u_{1j}$

   Mixed Model: $DISSEV_{ij} = \gamma_{00} + \gamma_{10}*RACE_{ij} + \gamma_{11}*IAT_j*RACE_{ij} + u_{0j} + u_{1j}*RACE_{ij} + r_{ij}$

3. *Does administrator implicit bias continue to adjust the relationship between race/PRS and discipline severity when controlling for student FRL eligibiligy, student grade, student infraction level, administrator experience, and administrator race? (Controlled Model)*

   Model: see above with addition of control variables as deemed appropriate by significance at the 0.05 level.

4. *Does administrator implicit bias predict the RDD?*

   Regression Model: $RDD_{ij} = \beta_{0j} + \beta_{1j}*(IAT_j) + r_{ij}$

Results

Preliminary Analyses

Response Bias

Due to low response rates, comparisons to determine whether participating and declining districts and/or administrators differed on study-related demographics and variables occurred prior to hypothesis testing. Groves and Peytcheva (2008) used a metareview of studies with various response rates to determine that response rates were not indicative of nonresponse bias and suggested comparing available data and paradata to screen for response bias. District level descriptive data is presented and compared in Table 6 and student level data in Table 7. Both participating and declining districts and students had highly variable data with likely outliers. As such, differences were only assumed for those variables with significant differences found via t-test in addition to a non-equal median. For districts, only the percentage of Black students in the district met these criteria, $t(15.9) = 2.91$, $p = .01$, median(declining) = 9%, median(participating) = 8%. Participating districts generally had lower percentages of Black students enrolled in school ($\bar{x} = 9, SD = 5.1$) than declining districts ($\bar{x} = 16, SD = 17.7$). Student data only met the criteria for grade in all discipline, $t(4159) = -14.68$, $p < .001$, median(declining) = 9, median(participating) = 10; and subjective discipline, $t(3984) = -13.97$, $p < .001$, median(declining) = 9, median(participating) = 10; subgroups, but not in objective discipline, $t(396) = -5.28$, $p < .001$, median(declining) = 9, median(participating) = 9.

Descriptive Statistics. Descriptive statistics were run on variables and paradata related to school/district enrollment. Table 8 provides demographic enrollment data for

participating schools and districts reduced from district level data to student level data to administrator level data and finally to student and administrator level data. The student and administrator level data describe the dataset used for hypothesis testing analyses.

Table 7 presents descriptive analyses of student level data used for hypothesis testing under the "Participating Administrator" heading. Although analyses did not include exclusionary discipline counts, extrapolation of exclusionary discipline counts to calculate RDD warranted descriptive analysis of this variable. Table 9 presents descriptive statistics for RDD and all other administrator level variables. Exclusionary discipline presented in Table 9 represents exclusionary discipline counts per administrator.

A Priori Assumption Testing

Data were screened *a priori* for collinearity with same level variables using Pearson's *r* correlations. Correlations presented in Tables 10a and 10b indicate no same-level multicollinearity of concern despite significant correlations between many variables because only variables entered in separate models are collinear at moderate or higher levels as determined by *r* values greater than 0.60 (Field, 2013). Values with strong correlations included discipline severity and exclusionary discipline which measure the same construct, and administrator race with counts of exclusionary discipline to students of Color. This final set of correlations, although unexpected, did not violate statistical testing assumptions because RDD models (based on exclusionary discipline) did not include administrator race.

Power

  Power analyses in multi-level modeling such as HLM are under-researched and not generally agreed upon (Reise, & Duan, 2003). Tools available for power analyses generally require equal cluster sizes (Snijders & Bosker, 1993) and employ simulation techniques. Due to varying cluster sizes, (All: $\bar{x} = 96, \sigma^2 = 13{,}309$; Subjective: $\bar{x} = 94, \sigma^2 = 12{,}018$; Objective: $\bar{x} = 16, \sigma^2 = 424$) power analyses using such software were impossible. Instead, the researcher made power evaluations based on sample sizes per the guidelines of Kreft (1996) suggesting that a sample with at least 30 clusters with at least 30 data points in each cluster would provide sufficient power, which occurred in all except for the objective discipline sample. Maas and Hox (2002) suggested at least 100 clusters with 10 data points in each when looking at random cross-level effects like that of the hypothesized model. When not possible, Maas and Hox (2002) noted that using only REML estimation and statistics with robust standard errors are best, and so those were used to provide more cautious and more powerful data interpretations.

  Power analyses for ordinary least squares (OLS) regression are not subject to the same level of controversy. Power analyses occurred *a priori* using G*Power 3.1.9.2 (Faul, Erdfelder, Buchner, & Lang, 2009). Using an alpha level of 0.05, power level of 0.80, and a null effect slope of 0, power estimates suggested a minimum sample size of 59 administrators for a larger effect size (slope = 0.35), 343 administrators for a medium effect size (slope = 0.15), and 19,617 administrators for a small effect size (slope = 0.02) per the regression effect sizes suggested by Faul, Buchner, & Lang (2009). These sample sizes were not obtained, so an additional *post hoc* power analysis was conducted using the same expectations. The actual sample size of 30 revealed an achieved power of 0.51

with a large effect size, 0.13 with a medium effect size, and 0.05 with a small effect size; which indicated no less than 49% chance of a false negative finding. Despite power levels highly subject to Type II error, results are presented in this paper with cautious interpretation.

Differences in Discipline Severity between Administrators (FUM)

To address the first research question regarding whether differences existed in student discipline severity at the administrator level and assess the need for multi-level analyses, fully unconditional models (e.g. empty models) were run for each set of student discipline data using the following formulae:

Level-1 Model: $DISSEV_{ij} = \beta_{0j} + r_{ij}$

Level-2 Model: $\beta_{0j} = \gamma_{00} + u_{0j}$

Mixed Model: $DISSEV_{ij} = \gamma_{00} + u_{0j} + r_{ij}$

All Discipline

A test of the FUM found a significant proportion of the variance in discipline severity occurred between administrators, Intra-class Correlation (ICC) = 0.31; $\chi^2(38) = 1224.32$, $p < .001$. Additionally, the design effect calculated based on an average of 94 students per administrator was 29.64, well above the 2.0 cutoff suggested by Muthén and Satorra (1995). Additional statistics related to this model are presented in Tables 11a and 12.

The ICC represents the proportion of variance, or differences, that occur at level 2 (between administrators) and is used to warrant use of HLM by determining if a reasonable amount of variance occurs outside the first level. Due to a lack of an agreed upon ICC cutoff and the ICC's inability to account for cluster sizes, the design effect is

also used to confirm a need for multi-level modeling Muthén and Satorra (1995). The design effect reflects the lost effectiveness by using cluster sample, and reflects how much larger or smaller of a sample is required when clustering.

Subjective Discipline

A test of the FUM found a significant proportion of the variance in subjective discipline severity occurred between administrators, ICC = 0.33; $\chi^2(38) = 1321.35$, $p < .001$. The design effect calculated based on an average of 90 students per administrator was 30.64, again above the 2.00 cutoff. Additional statistics related to this model are presented in Tables 11b and 12.

Objective Discipline

A test of the FUM found a significant proportion of the variance in objective discipline severity occurred between administrators, ICC = 0.22; $\chi^2(32) = 98.91$, $p < .001$. The design effect calculated based on an average of 16 students per administrator was 4.26, which although lower than the other design effects, remained above the 2.00 cutoff. Additional statistics related to this model are presented in Tables 11c and 12.

Summary

The FUM statistics for each set of disciplinary data indicated significant proportions of the variances in discipline severities occurred between administrators with 31% overall, 33% in subjective discipline, and 22% in objective discipline.

<div align="center">

Hypothesized Model

</div>

Once differences in discipline severity at the administrator level were confirmed, hierarchical analyses were deemed necessary. The second research question focused on the central hypothesis linking administrator implicit bias to differences in the relationship

between student race/PRS and discipline severity. The formulae for this hypothesized

model follows:

Level-1 Model: $DISSEV_{ij} = \beta_{0j} + \beta_{1j}*(RACE_{ij}) + r_{ij}$

Level-2 Model: $\beta_{0j} = \gamma_{00} + u_{0j}$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}*(IAT_j) + u_{1j}$$

Mixed Model: $DISSEV_{ij} = \gamma_{00} + \gamma_{10}*RACE_{ij} + \gamma_{11}*IAT_j*RACE_{ij} + u_{0j} +$

$u_{1j}*RACE_{ij} + r_{ij}$

No variables were centered when added to the model because zero acted as a comparison

group (White Students) for the race/PRS variable and a true zero existed on the measure

of implicit bias.

All Discipline

The model converged after 366 iterations with inclusion of 3,432 student level

records, and 39 administrator-level records (variance components and reliability

estimates included only 35 administrators due to insufficient data). The fixed effect (e.g.,

intercept) of race/PRS was a significant positive predictor (i.e., as skin color became

darker discipline severity became more severe) of mean discipline severity, $\beta = -0.09$,

$t(37) = 3.56$, $p = .001$. Similarly, the fixed effect of IAT score was a significant negative

predictor (i.e., as bias score became more pro-White discipline severity became less

severe) of the slope between student race/PRS and discipline severity, $\beta = -0.11$, $t(37) = -$

2.51, $p = .017$. The intercept slope (e.g., random effect, level one variance) was

significant indicating enough variance for additional predictors of mean discipline

severity, $\chi^2(34) = 485.81$, $p < .001$. The race/PRS slope (e.g., random effect of race/PRS,

variance in the race/PRS to discipline slope) was not significant indicating there was

insufficient variance remaining for additional predictors of the student race/PRS to discipline severity slope, $\chi^2(33) = 27.11$, $p > .05$. The expected discipline severity for a White student receiving discipline from an administrator regardless of IAT score was 2.91, but for a Black student from an administrator with no implicit bias (IAT score = 0) was 3.08 or from an administrator with a moderate preference for White students (IAT score = -.5) was 3.19. Additional model statistics are presented in Table 11a.

Pseudo R^2 values calculated for this model are displayed in Table 12 and indicated that student race/PRS and administrator IAT scores accounted for 4% of the variance in discipline severity between students (within administrators) and 14% of the variability between administrators. Administrator IAT scores accounted for 87% of the variance in discipline severity by student race/PRS.

Subjective Discipline

The model converged after 422 iterations with inclusion of 3278 student level records, and 39 administrator-level records (variance components and reliability estimates included only 35 administrators due to insufficient data). The fixed effect of race/PRS was a significant positive predictor of mean subjective discipline severity, $\beta = 0.08$, $t(37) = 3.21$, $p = .003$. Similarly, the fixed effect of IAT score was a significant negative predictor of the slope between student race/PRS and subjective discipline severity, $\beta = -0.09$, $t(37) = -0.09$, $p = .027$. The intercept slope was significant indicating enough variance for additional predictors of mean subjective discipline severity, $\chi^2(34) = 524.44$, $p < .001$. The race/PRS slope was not significant indicating there was insufficient variance remaining for additional predictors of the student race/PRS to subjective discipline severity slope, $\chi^2(33) = 29.54$, $p > .05$. The expected discipline severity for a

White student receiving discipline from an administrator regardless of IAT score was 2.91, but for a Black student from an administrator with no implicit bias (IAT score=0) was 2.96 or from an administrator with a moderate preference for White students (IAT score = -.5) was 3.04. Additional model statistics are presented in Table 11b.

Pseudo R^2 values calculated for this model are displayed in Table 12 and indicated that student race/PRS and administrator IAT scores accounted for 3% of the variance in subjective discipline severity between students (within administrators) and 9% of the variability between administrators. Administrator IAT scores accounted for 89% of the variance in subjective discipline severity by student race/PRS.

Objective Discipline

The model converged after 356 iterations with inclusion of 263 student level records, and 32 administrator-level records (variance components and reliability estimates included only 15 administrators due to insufficient data). The fixed effect of race/PRS was a significant positive predictor of mean objective discipline severity, $\beta = 0.35$, $t(30) = 3.15$, $p = .004$. The fixed effect of IAT score was not a significant predictor of the slope between student race/PRS and objective discipline severity, $t(30) = 1.78$, $p = .085$. The intercept slope was significant indicating enough variance for additional predictors of mean objective discipline severity, $\chi^2(14) = 43.67$, $p < .001$. The race/PRS slope was not significant indicating there was insufficient variance remaining for additional predictors of the student race/PRS to objective discipline severity slope, $\chi^2(13) = 21.17$, $p = .069$. The expected discipline severity for a White student receiving discipline was 4.26, but 4.96 for a Black student. Additional model statistics are presented in Table 11c.

Pseudo R^2 values calculated for this model are displayed in Table 12 and indicated that student race/PRS and administrator IAT scores accounted for 1% of the variance in objective discipline severity between students (within administrators). A negative R^2 value when using both the Snijders & Bosker (1994, 1999) and Kreft & De Leeuw (1998)/ Singer (1998) methods resulted in no viable measure of level 2 variance.

Summary

Variance accounted for at the student level was similar for subjective and overall discipline ($R^2_{1A} - R^2_{1S} = 0.04$), but dissimilar for objective discipline ($\frac{(R^2_{1A} + R^2_{1S})}{2} - R^2_{1O} = 0.12$). At the administrator level, variance in mean discipline severity followed the same pattern with similar variance accounted for in models of overall and subjective discipline ($R^2_{2A} - R^2_{2S} = 0.03$) and dissimilar variance accounted for in the objective discipline model discipline ($\frac{(R^2_{2A} + R^2_{2S})}{2} - R^2_{2O} = 0.62$). A pseudo R^2 for the student race/PRS – objective discipline severity slope was inappropriate due to non-significance, but overall and subjective discipline continued to account for similar levels of variance ($R^2_{2sS-A} - R^2_{2sS} = 0.02$).

Controlled Model

After the hypothesized model included a significant relationship, control variables were added to the model to assess whether the hypothesized effects remained despite control for other possibly related indicators of discipline severity. Because the objective discipline model did not confirm implicit bias as a predictor of the relationship between student race/PRS and objective discipline severity, controls were added to this model

without the predictor of IAT score. The most detailed formulae possible for the controlled model follows:

Level-1 Model: $DISSEV_{ij} = \beta_{0j} + \beta_{1j}*(FRL_{ij}) + \beta_{2j}*(RACE_{ij}) + \beta_{3j}*(GRADE_{ij})$

$$+ \beta_{4j}*(INFLVL_{ij}) + r_{ij}$$

Level-2 Model: $\beta_{0j} = \gamma_{00} + \gamma_{01}*(ADEXP_j) + \gamma_{02}*(ADRACE_j) + u_{0j}$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}*(IAT_j) + u_{2j}$$

$$\beta_{3j} = \gamma_{30}$$

$$\beta_{4j} = \gamma_{40}$$

Mixed Model: $DISSEV_{ij} = \gamma_{00} + \gamma_{01}*ADEXP_j + \gamma_{02}*ADRACE_j + \gamma_{10}*FRL_{ij} +$

$$\gamma_{20}*RACE_{ij} + \gamma_{21}*IAT_j*RACE_{ij} + \gamma_{30}*GRADE_{ij} +$$

$$\gamma_{40}*INFLVL_{ij} + u_{0j} + u_{2j}*RACE_{ij} + r_{ij}$$

All control variables were added as fixed effects due to expected consistency in the effect of each variable across administrators despite different values between students. In short, the expected effects of a control variable on discipline severity did not depend on the administrator responsible for the discipline. All variables were added without centering except for infraction level (INFLVL) based on the absence or presence of meaningful zero values. Reference values for FRL eligibility, grade, administrator race (ADRACE), and administrator experience (ADEXP) were paid lunch, kindergarten, White, and less than one year, respectively. Infraction level was centered around the group mean despite being fixed due to a large level of variability between administrators' mean infraction levels (SD = 0.58; Range = 2.33; Scale: 1-6). As such, result interpretations of change in

infraction level occurred relative to the mean of each administrator rather than the overall mean across administrators.

All Discipline

The model converged after 478 iterations with inclusion of 3430 student level records, and 39 administrator-level records (variance components and reliability estimates included only 35 administrators due to insufficient data). The final model's formulae were:

Level-1 Model: $DISSEV_{ij} = \beta_{0j} + \beta_{1j}*(FRL_{ij}) + \beta_{2j}*(RACE_{ij}) + \beta_{4j}*(INFLVL_{ij}) + r_{ij}$

Level-2 Model: $\beta_{0j} = \gamma_{00} + u_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20} + \gamma_{21}*(IAT_j) + u_{2j}$

$\beta_{3j} = \gamma_{30}$

Mixed Model: $DISSEV_{ij} = \gamma_{00} + \gamma_{10}*FRL_{ij} + \gamma_{20}*RACE_{ij} + \gamma_{21}*IAT_j*RACE_{ij} +$

$\gamma_{40}*INFLVL_{ij} + u_{0j} + u_{2j}*RACE_{ij} + r_{ij}$

As controls were added to the model, grade, administrator race, and administrator experience were insignificant, Grade: $t(3349) = 0.12$, $p = .901$, ADRACE: $t(37) = 1.29$, $p = .205$, ADEXP: $t(37) = -0.04$, $p = .971$. The fixed control variables of infraction level and FRL eligibility were significant positive predictors of mean discipline severity (i.e., as infractions became more severe or students became eligible for free lunch, discipline severity became more severe), INFLVL: $\beta = 0.13$, $t(3350) = 13.17$, $p < .001$; FRL: $\beta = 0.41$, $t(3350) = 5.64$, $p < .001$. After controlling for these variables, the fixed effect of race/PRS was no longer a significant predictor of mean discipline severity, $t(37) = 1.50$, $p = .143$. Similarly, administrator IAT score was no longer a significant predictor of the

88

slope between student race/PRS and discipline severity, $t(37) = -2.03$, $p = .05$. The intercept slope was significant indicating enough variance for additional predictors of mean discipline severity, $\chi^2(34) = 555.43$, $p < .001$; but the race/PRS slope was not significant indicating there was insufficient variance remaining for additional predictors of the student race/PRS to discipline severity slope, $\chi^2(33) = 32.71$, $p > .05$. Additional model statistics are presented in Table 11a.

Pseudo R^2 values calculated for this model are displayed in Table 12 and indicated that the model accounted for 22% of the variance in discipline severity between students (within administrators) and 21% of the variability between administrators.

Subjective Discipline

The model converged after 347 iterations with inclusion of 3276 student level records, and 39 administrator-level records (variance components and reliability estimates included only 35 administrators due to insufficient data). The final model's formulae were:

Level-1 Model: $DISSEV_{ij} = \beta_{0j} + \beta_{1j}*(FRL_{ij}) + \beta_{2j}*(RACE_{ij}) + \beta_{4j}*(INFLVL_{ij}) + r_{ij}$

Level-2 Model: $\beta_{0j} = \gamma_{00} + u_{0j}$

$\beta_{1j} = \gamma_{10}$

$\beta_{2j} = \gamma_{20} + \gamma_{21}*(IAT_j) + u_{2j}$

$\beta_{3j} = \gamma_{30}$

Mixed Model: $DISSEV_{ij} = \gamma_{00} + \gamma_{10}*FRL_{ij} + \gamma_{20}*RACE_{ij} + \gamma_{21}*IAT_j*RACE_{ij} +$

$\gamma_{40}*INFLVL_{ij} + u_{0j} + u_{2j}*RACE_{ij} + r_{ij}$

In adding controls to the model student grade, administrator race, and administrator experience were insignificant predictors of mean subjective discipline severity, Grade:

$t(3195) = -0.02$, $p = .986$, ADRACE: $t(37) = 1.00$, $p = .326$, ADEXP: $t(37) = 0.11$, $p =$ .915. The fixed control variables of infraction level and FRL eligibility were significant positive predictors of mean subjective discipline severity, INFLVL: $\beta = 0.13$, $t(3196) =$ 7.97, $p < .001$; FRL: $\beta = 0.42$, $t(3196) = 5.72$, $p < .001$. After controlling for these variables, the fixed effect of race/PRS remained a significant predictor of mean subjective discipline severity, $\beta = 0.06$, $t(37) = 2.09$, $p = .043$. Similarly, administrator IAT score remained a significant predictor of the slope between student race/PRS and subjective discipline severity, $\beta = -0.08$, $t(37) = -2.18$, $p = .036$. The random effect associated with the intercept was significant indicating enough variance for additional predictors of mean subjective discipline severity, $\chi^2(34) = 478.31$, $p < .001$; but the race/PRS slope was not significant insufficient remaining variance for any additional predictors of the student race/PRS to subjective discipline severity slope, $\chi^2(33) = 35.45$, $p = .351$. The expected discipline severity for a White student, ineligible for FRL, with an infraction level equal to the mean infraction level of his or her school receiving discipline from an administrator regardless of IAT score was 2.65, but for a Black student under the same conditions receiving discipline from an administrator with no implicit bias (IAT score=0) was 2.76 or from an administrator with a moderate preference for White students (IAT score = -.5) was 2.86. Additional model statistics are presented in Table 11b.

Pseudo R^2 values calculated for this model are displayed in Table 12 and indicated that the final model accounted for 13% of the variance in subjective discipline severity between students (within administrators) and 21% of the variability between administrators. Furthermore, the final model accounted for 89% of the variability in the

slope between student race/PRS and subjective discipline severity, which is equal to the accounted variability before controls.

Objective Discipline

The model converged after 1772 iterations with inclusion of 263 student level records, and 32 administrator-level records (variance components and reliability estimates included only 15 administrators due to insufficient data). The final model's formulae were:

Level-1 Model: $DISSEV_{ij} = \beta_{0j} + \beta_{1j}*(RACE_{ij}) + \beta_{2j}*(INFLVL_{ij}) + r_{ij}$

Level-2 Model: $\beta_{0j} = \gamma_{00} + u_{0j}$

$\beta_{1j} = \gamma_{10} + u_{1j}$

$\beta_{2j} = \gamma_{20}$

Mixed Model: $DISSEV_{ij} = \gamma_{00} + \gamma_{10}*RACE_{ij} + \gamma_{20}*INFLVL_{ij} + u_{0j} + u_{1j}*RACE_{ij} + r_{ij}$

In adding controls to the model student FRL eligibility, student grade, administrator race, and administrator experience were insignificant predictors of mean objective discipline severity, FRL: $t(197) = 0.24$, $p = .813$, GRADE: $t(197) = 0.76$, $p = .434$, ADRACE: $t(30) = 0.27$, $p = .793$, ADEXP: $t(30) = -0.36$, $p = .723$. The fixed control variable of infraction was a significant positive predictor of mean objective discipline severity, $\beta = 0.27$, $t(198) = 3.50$, $p < .001$. After controlling for these variables, the fixed effect of race/PRS remained a significant positive predictor of mean objective discipline severity, $\beta = 0.19$, $t(31) = 2.24$, $p = .032$. The random effect associated with the intercept was significant indicating enough variance for additional predictors of mean objective discipline severity, $\chi^2(14) = 46.50$, $p < .001$; but the race/PRS slope was not significant insufficient remaining variance for any additional predictors of the student race/PRS to objective

discipline severity slope, $\chi^2(14) = 22.61$, $p = .067$. The expected discipline severity for a White student, with an infraction level equal to the mean infraction level was 4.28, but for was 4.67 for a Black student under the same conditions. Additional model statistics are presented in Table 11c.

Pseudo R^2 values calculated for this model are displayed in Table 12 and indicated that the final model accounted for 11% of the variance in objective discipline severity between students (within administrators) and 71% of the variability between administrators.

Summary

After inclusion of control variables typically associated with differences in discipline severity and/or administrator behavior, only the hypothesized model for subjective discipline remained significant and accounted for 89% of the variability in the relationship between student race/PRS and subjective discipline severity. Graphical depictions based on significant predictors in final HLM models are depicted in Figure 4. The overall discipline severity model accounted for the most variance between students (within administrators) at 22% accounted, and the subjective and objective discipline models accounted for similar amounts of variance at this level, $R^2_{1S} - R^2_{1O} = 0.02$. At the administrator level, the objective discipline model accounted for a much larger amount of the variance between administrators than the other models with 71% accounted for as comparted to 21% at both the overall and subjective levels. All models demonstrated a positive association between race/PRS and discipline severity wherein as race/PRS moved towards a student more associated with a skin tone of Color, the severity of discipline increased as demonstrated in Figure 4.

92

Racial Discipline Discrepancy and Implicit Bias

Simple linear regression indicated that implicit bias was not a significant predictor of RDD for overall discipline, $F(1, 29) = 0.01$, $p = .921$. Due to the significant findings of subjective-only discipline in the HLM models, a repetition of the RDD regression analysis occurred for subjective-only discipline. Simple linear regression again indicated that implicit bias was not a significant predictor of RDD for subjective discipline, $F(1, 29) = 0.30$, $p = .588$. Graphs of the data for each model are presented in Figure 5.

Assumption Tests

Overall Discipline Severity Model

The final model was assessed for assumption violations including homogeneity and homoscedasticity of variance and normality of residuals. The significance of several effects confirmed the assumption of linearity for significant variables in the final model. To assess linearity in non-significant variables, fixed effects (student grade, administrator race, and administrator experience) were plotted against discipline severity and the random effect (IAT score) was separated by high and low scores on a graph of the relationship between student race/PRS and overall discipline as presented in Figure 6. No non-linear relationships were apparent. The assumption of homogeneity of level one variance was violated, $\chi2(34) = 267.88$, $p = .000$, but an expected skew towards less severe discipline is a likely driver of this violation. Level one residuals were within the +2 range for a normally distributed data (Field, 2013) with a skewness of 0.26 (SE=.04) and a kurtosis of -0.99 (SE=.08). The assumption of homoscedasticity of level 1 residuals was met as seen by the lack of patterning in Figure 7. At level 2, residuals again met the assumptions of normality with a skewness of -0.10 (SE=.38) and kurtosis of -0.34

(SE=.74) for the intercept residuals and a skewness of -0.10 (SE=.38) and kurtosis of -0.37 (SE=.74) for the student race/PRS slope residuals. Homoscedasticity was uninterpretable for the intercept due to only one fit value as seen in Figure 8, but homoscedasticity of the student race/PRS slope residuals appeared present as demonstrated in Figure 9. The assumption of level two multivariate normality was present as shown by the near one-to-one linear relationship (usually seen as a 45-degree angle) in Figure 10.

Subjective Discipline Severity Model

The final model was assessed for assumption violations beginning with the significance of several effects confirming the assumption of linearity for significant variables in the final model. Non-significant variables (student grade, administrator race, and administrator experience) were assessed graphically as seen in Figure 11. No non-linear relationships were apparent. The assumption of homogeneity of level one variance was again violated, $\chi^2(34) = 285.16$, $p = .000$. Level one residuals were normally distributed with a skewness of 0.38 (SE=.043) and a kurtosis of -0.08 (SE=.084). The assumption of homoscedasticity of level 1 residuals was met as seen by the lack of patterning in Figure 7. At level 2, residuals again met the assumptions of normality with a skewness of 0.13 (SE=.38) and kurtosis of -0.37 (SE=.74) for the intercept residuals and a skewness of 0.13 (SE=.38) and kurtosis of -0.39 (SE=.74) for the student race slope residuals. Homoscedasticity was again uninterpretable for the intercept due to only one fit value as seen in Figure 8, but homoscedasticity of the student race/PRS slope residuals was met as demonstrated in Figure 9. The assumption of level two multivariate normality seemed met as shown by the near one-to-one linear relationship (usually seen as a 45-

94

degree angle) in Figure 10 except for one outlier-like value. This value was evaluated for unlikely attributes, but none were found.

Objective Discipline Severity Model

The final model was assessed for assumption violations including linearity as confirmed by the significance of several effects in the final model. Non-significant variables (student FRL eligibility, student grade, administrator race, and administrator experience) and the random effect of IAT score on the student race/PRS-discipline severity relationship were assessed graphically as presented in as seen in Figure 12. No non-linear relationships were apparent. The assumption of homogeneity of level one variance was again violated, $\chi^2(12) = 137.08$, $p = .000$. Level one residuals were not normally distributed with a normal skewness of -1.13 (SE=.15) and a non-normal kurtosis of 2.08 (SE=.30). The assumption of homoscedasticity of level 1 residuals was likely met as seen by the lack of overall patterning despite some clustering in the upper right portion as seen in Figure 7. At level 2, residuals again violated the assumptions of normality with a normal skewness of -1.83 (SE=.41) and non-normal kurtosis of 3.36 (SE=.81) for the intercept residuals and a normal skewness of 1.82 (SE=.38) and non-normal kurtosis of 3.25 (SE=.81) for the student race/PRS slope residuals. Homoscedasticity was once again uninterpretable for the intercept due to only one fit value as seen in Figure 8, with homoscedasticity of the student race/PRS slope residuals following the same pattern as displayed in Figure 9. Figure 10 displays a moderate one-to-one linear relationship indicating a likely adherence to the assumption of level two multivariate normality.

95

Racial Discipline Gap

The use of continuous predictors with linear relationships for all RDD regression analyses were confirmed graphically in Figure 5. This figure also displays the presence of outliers in the RDD for all discipline; however, no reasonable reason to remove these values from analyses existed. The assumption of independence of observations was violated with a possible positive auto-correlation, DW(all)= 0.50, DW(subjective)=0.63, dL(29,1)= 1.12, $p < .010$. Figure 5 displays no pattern around the line of best fit for either discipline type indicating homoscedasticity. A histogram and Normal P-P plot display a violation of the assumption of normally distributed residuals for overall discipline but not for subjective discipline in Figure 13.

Summary

Models of discipline severity demonstrated that in only subjective discipline, the effect of administrator implicit bias on the relationship between student race/PRS and discipline severity persisted after controlling for infraction level and student FRL eligibility. Furthermore, objective discipline severity differed only by infraction level, and was unaffected by student race/PRS, FRL eligibility, or grade and by administrator race and experience. RDD was not predicted by implicit bias even when separated into subjective-only discipline.

CHAPTER 5:

Discussion

The current study attempted to provide a clearer understanding of the relationship between school administrators' implicit bias and the school racial discipline gap. The researcher used the IAT to measure administrators' implicit bias as a predictor of the subjective, objective, and overall discipline severity to student race/PRS relationship. Additionally, the study assessed the predictive value of administrator implicit bias on RDD, a determinant of the racial discipline gap based on the discipline associated with one administrator. In these ways, the researcher worked towards understanding whether implicit bias influenced the discipline-related decision-making of PA K-12 school administrators. This chapter discusses these finding by first interpreting the statistical data presented in Chapter 4, and then assessing the significance and contributions of findings and non-findings. Subsequently, limitations of the study are discussed to contextualize next steps and future research directions. Recommendations for practice are presented before the full dissertation is summarized.

Inconsistent Discipline Severity

Research Question 1 explored whether discipline severity differed by administrators and determined if HLM was necessary. Regardless of the level of decision-making involved in the infraction, a sizeable proportion of the variance in discipline severity occurred between administrators. Administrator differences accounted for 33% of differences in subjective discipline decisions and 22% of differences in objective discipline decisions, suggesting administrators made decisions regarding discipline differently. Even when evaluating discipline overall, 31% of differences

97

occurred between administrators. Although seemingly suggesting administrators assign disciplinary actions in their own ways regardless of the level of regulation involved in the discipline decision-making process, other possible explanations remain more prospective. Different school districts and often schools set different discipline policies through codes of conduct, discipline scaling plans, and other methods. With the inability to account for the nesting of administrators in schools and/or school districts, this model cannot determine if differences in objective disciplinary decisions were truly differences by administrator or different school/district policies. For subjective discipline, schools and districts typically neglect to provide associated policies or provide only loose policies with a range of appropriate disciplinary outcomes. Hence, the results indicate subjective disciplinary decisions differ by administrators in a manner subject to bias. Overall discipline findings provided unclear information regarding whether differences are likely to occur at the school/district or just the administrator level. Together, all models of discipline suggested additional exploration and continued hierarchical analyses.

Implicit Bias and Racial Discipline Severity Differences

Research Question 2 evaluated the hypothesized model considering whether administrators' implicit bias could predict differences in the relationship between student race/PRS and discipline severity. Administrators with higher levels of implicit bias chose more severe disciplinary actions for students of Color than administrators with lower levels of implicit bias for overall and subjective discipline, but not objective discipline. For objective discipline, students received more severe discipline if they identified as of Color consistently despite administrator implicit bias. These race/PRS-based differences by administrator potentially related to one of the control variables or the school/district

98

context where the administrator was situated. Although a potential lack of power due to considerably less objective discipline cases might share responsibility for insignificant findings, the presence of a significant race/PRS effect suggests enough power to find at least large effect sizes. If implicit bias legitimately failed to predict differences in the student race/PRS-student discipline severity relationship, then choice-based decisions, as expected, are more vulnerable to the effects of implicit bias. Hypothetically, objective decisions in discipline should not allow for bias due to the pre-determined nature of the outcomes based on the infractions; however, if policies are themselves biased one might expect objective discipline differences between administrators as well unless a third level, modeling school/district effects, results differently.

The confirmation of the hypothesized effect of implicit bias on the race/PRS-discipline severity relationship in subjective and objective discipline presents interesting information as well. In both subjective and overall discipline, IAT scores accounted for a large portion of the variance in the race/PRS-discipline severity slope (89% and 87%, respectively). However, the subjective discipline model only explained 3% of the between student and 9% of the between administrator differences, and the overall model only explained 4% of the between student and 14% of the between administrators. Despite lower values, the variance explained at these levels is not inadequate based on the highly variant behavior of human beings. Significant findings confirmed that a relationship existed, leaving lower explained variability values possibly related to the presence of outliers and/or widely spread data values.

The level of implicit bias associated with each administrator significantly predicted changes in both overall and subjective, but not objective, discipline based on

student race/PRS. The results imply that when discipline involved less-controlled decision-making, implicit racial biases threatened the equity of disciplinary actions assigned by school administrators. Subjective decisions in discipline depended on implicit bias even more than overall decisions. As such, the findings warranted considerations of whether expected correlates of discipline severity could remove the significance of implicit bias as a predictor of the race/PRS-discipline severity slope or even the racial differences in discipline severity all together.

The Perseverance of Implicit Bias by Decision Type

The final set of models of discipline severity added the predictors of infraction level (severity of behavior leading to discipline), FRL eligibility (indicative of SES), student grade, administrator experience, and administrator race. These predictors have been linked to differences in discipline severity previously (see Skiba, et. al, 2014) and presented the potential to remove the significance of the effects of implicit bias. Grade, administrator experience, and administrator race were insignificant predictors of discipline severity in all models. This suggested a possible dissimilarity between PA and other states and problematic sample-imbalances reflecting race and grade. Most students sampled were in high school (grades 9-12), followed by a large portion of students in middle school (grades 5-8). Very few students represented elementary grades and many of those students committed higher level infractions resulting in high level discipline. Furthermore, most elementary students received objective discipline further removing this group from the samples. Only three of the 39 administrators sampled were "of Color" with the remainder identifying as White. This likely placed a heavy weight on the responses of the administrators of Color and possibly reduced the likelihood of seeing

100

administrator race-based effects if they were present. Administrators sampled represented a diverse range of experience indicating a more reliable insignificant result.

As expected, infraction level and race/PRS were significant predictors of discipline severity in all samples. Student FRL eligibility was a significant predictor of overall and subjective discipline severity, but not objective discipline severity. Potentially, social factors such as FRL eligibility and grade affect objective discipline less, adding confidence to the significance and insignificance for different objective discipline models. When adding these significant controls to each model, only the subjective discipline model retained significance for both race/PRS and implicit bias. The overall discipline model was no longer significantly different based on race/PRS or implicit bias, and the objective discipline model (run only with race/PRS as a predictor only) confirmed the persistence of the race/PRS-based finding. Summarizing, the predictive effects of infraction level and FRL eligibility were strong enough to remove the differences attributed to race/PRS for overall discipline. Nevertheless, the effect of implicit bias on the race/PRS-discipline severity relationship seen in subjective discipline remained present despite control for FRL eligibility and infraction level.

In overall discipline, students who received free lunch had more severe discipline based on more severe infractions with infraction level and FRL eligibility accounting for 22% and 21% of the variability between students and between administrators, respectively. When compared to the hypothesized model, infraction level and FRL eligibility were solely responsible for 18% of the differences between students and 6% of the between-administrator variability. The lower proportion of variance accounted for between administrators suggests that infraction level and FRL eligibility work in a more

global manner rather than by driving differences in discipline severity allocation between administrators.

For objective discipline, 11% of the differences in discipline severity occurred between students while 10% of differences between students were solely attributed to infraction level. Between administrators, infraction level explained 71% of differences. Because administrators typically follow policy for objective decisions, the large predictive value of infraction level between administrators likely stems from differences in policy between administrators for which a school/district level in the model could account. In short, students experienced more severe discipline in response to objective-type infractions based on both race/PRS and infraction level.

Subjective disciplinary decisions remained vulnerable to the effects of implicit bias through student race/PRS despite controls for infraction level and FRL eligibility indicating that the effect of implicit bias on the race/PRS to discipline severity relationship was not simply an artifact of SES or differences in behavior severity by race/PRS. In fact, infraction level and FRL eligibility added no additional predictive value to the slope between race/PRS and subjective discipline severity. Infraction level and FRL eligibility accounted solely for 10% of the differences between students and 12% of the differences between administrators. Students of Color experienced more severe discipline severity per administrators' implicit bias despite the students' SES or infraction level. Interestingly, SES seemed to lessen the effects of implicit bias on the relationship between race/PRS and discipline severity as seen in Figure 4. Based on the association between SES and student race/PRS, this is not surprising, but the near inversion of the relationship signifies a need for further study.

102

The Racial Discipline Gap by Administrator

To determine if administrator implicit bias was directly involved in the racial school discipline gap, Research Question 4 investigated whether implicit bias regressed with the RDD, or gap between the proportion of students of Color in a school and the proportion of students of Color receiving exclusionary discipline from a given administrator in that school. Results were split into overall discipline and subjective discipline, but no significant findings were found. The sample size in this analysis was under that required for sufficient power, making the possibility of Type II error high. Similarly, the limited dataset failed to meet several of the assumptions of simple linear regression creating results that might not accurately reflect the data. As such, interpretation of insignificance in the relationship between administrator implicit bias and the RDD presents only a possible interpretation of the findings under the assumption of no Type II error.

A lack of a connection between administrator implicit bias and the localized RDD indicated the school racial discipline gap may not stem from implicit bias at the administrator level. Despite findings demonstrating administrator implicit bias contributed to differences in discipline severity by race/PRS, the extent of these differences possibly falls short of translation to the RDD or does not extend into inclusive versus exclusive discipline. Research has already confirmed that teachers over-refer students of Color for administrative discipline despite similar behaviors (Finn & Servos, 2015; Nicholson-Crotty, Birchmeir, & Valentine, 2008; Skiba, Shure, & Wilson, 2012; Wright, 2015), creating the possibility that this gap is truly a teacher-level challenge. It is also possible that calculation of a localized RDD was not reflective of the school racial

discipline gap overall. Current estimates of the racial discipline gap do not typically separate into subjective and objective discipline decisions or consider gaps on a school-by-school basis. As was the case in the data analyses presented in this study for objective/subjective versus overall discipline, discipline separated by schools may not approximate the same patterns as when considered overall. To the researcher's knowledge, the current study is the first to use RDD as a localized model of the greater school racial discipline gap creating questionable reliability and validity for this measure. Nonetheless, the findings suggest that implicit bias is not an effective predictor of RDD at the administrator level and fails to account for administrator-level discipline gaps associated with student race/PRS.

Significance and Contributions

Taken together, the findings of this study add to the current literature on racial implicit bias in schools by adding administrator implicit bias to considerations regarding inequitable discipline by race/PRS. Previous studies have focused only on teachers when trying to understand the role of implicit bias in school discipline (Gershenson, Holt, & Papageorge, 2015; Gilliam, Maupin, Reyes, Accavitti, & Shic, 2016; Glock & Karbach, 2015; Glock, Kneer, & Kovacs, 2013; Okonofua, & Eberbardt, 2015; van den Bergh, Denessen, Hornstra, & Holland, 2010; Wright, 2015), but this study introduced the implicit bias of the school administrator as an additional source of inequity in school discipline. Not only is this significant because it continues to answer questions regarding the sources of racial inequities in discipline, but it all provides a framework for reduction of such inequities through implicit bias mitigation strategies. These strategies are beginning to gain clarification as researchers work to test different practices. Some of

104

these practices such as counter-example exposure, mindfulness, and data review have existed in other areas related to education for quite some time while newer methods such as implicit bias measurement, recognition and response, and structured decision-making continue to gain popularity (Skiba, Mediratta, & Rausch, 2016). The results of this study can help inform school leaders on what skills and training administrators involved in disciplinary decision-making require for successful reduction of inequitable school discipline.

The persistence of the effect of implicit bias on the race/PRS-discipline relationship for only subjective disciplinary decisions contributes to the field by first highlighting the need to partition discipline into objective and subjective, second clarifying the critical role of school policy on mitigating bias, and third confirming subjective discipline as highly vulnerable to implicit biases despite schools' efforts to work towards equitable discipline. Studies of discipline where decision-making is focal must delineate between subjective and objective discipline to truly describe issues. Furthermore, delineation in such a manner at the state and federal levels of data collection has the potential to decrease the effects of bias on questionably subjective discipline while creating clarity in policy and law based discipline versus decision related discipline. This study demonstrated that in objective discipline, administrators' implicit bias did not affect the race/PRS-discipline relationship. As such, creating a clearer division between objective and subjective discipline could increase the quantity of decisions committed to objective disciplinary patterns. With reporting of discipline type, schools might feel obliged to further clarify policies, and legal entities might begin to clarify laws to develop more equitable policies considering any racially discrepant

105

findings pertaining to objective-only discipline. Finally, confirming the vulnerability of subjective discipline to bias creates an awareness of the risk involved with subjective discipline and may encourage schools to work towards bias mitigation.

This study also introduced the concept of a localized RDD as a potential single-subject measure for related discipline gaps. Although results suggested that this is not a valid model of the overall school discipline gap, schools may use this measure to determine if individual administrators or the school as a whole are contributing most to discipline disparities by race/PRS. Furthermore, the measure provides a simple calculation for other disparities in school populations such as gender representation in STEM courses, LGBTQ+ representation in athletics, or even representation in advanced placement courses by SES. Using Reschley's (1997) determinant of discrepancy with the model of the RDD calculation employed here provides an effective discrepancy measurement for a variety of situations of concern to school leaders and educational researchers alike.

Limitations

Despite interesting findings, the current study suffers from a multitude of limitations rooting from threats of error, sampling issues tied to response rates, study design challenges, and limitations inherent in research on sensitive topics. When studying issues that are often considered private or threatening, such as oppression and privilege, individuals often question methodology (Nance, 2016; Tetlock & Mitchell, 2009), hesitate to participate, and filter responses (Hatchett, & Schuman, 1976). As such, some individuals choose to disbelieve the IAT is a measure of implicit bias or disregard

implicit bias as a phenomenon all together (Jost, Rudman, Blair, Carney, Dasgupta, Glaser, & Hardin, 2009; Tetlock & Mitchell, 2009). Jost and his team suggested:

> Resistance is all the more likely when social scientific discoveries seem to challenge long cherished personal or cultural assumptions, such as the relatively hopeful messages that (a) human thought and behavior are largely under the control of individual will and consciousness, and (b) racial prejudice in Western societies (especially the U.S.) is a thing of the past. (Jost, et al., 2009, p. 41)

Based on Jost's assessment, implicit bias presents a challenge to individuals' confidence that behaviors are conscious and chosen and that racism is no longer an issue of concern. Hence, findings from this study may prove hard to accept—presenting a challenge for incorporation into discipline literature and school practice. Furthermore, schools could filter shared data to lessen actual racial discipline gaps and/or administrators could alter self-reported data to present more positive reflections of themselves. Use of the IAT aimed to lessen such biases in self-reported data by using a validated measure focused on response times rather than thought-out answers. Nonetheless, this study of concepts linked to oppression created a high risk for potential participants, evidenced by the low response rate.

Response Rate

As noted earlier, only 5% of invited administrators and 11% of invited school districts chose to participate in the study. The initial study design included three levels with data at the student discipline level, administrator level, and district level to account for separate pools of variance in each group. Unfortunately, the study lacked the minimum of 30 districts necessary to run a three-level model with a reasonable level of

power. Fortunately, 39 administrators participated allowing for a two-level model. Unfortunately, the presence of only 30 administrators appropriate for inclusion in the RDD analysis provided too small a sample size for reasonable power.

Another issue with participating districts was the extremely low level of diversity in participating administrators; with only three out of 41 administrators identifying as "of Color" reasonable controls for administrator race were likely impossible due to high sample homogeneity. Similarly, student discipline occurred mostly at the high school and moderately at the middle school level. This under-representation of elementary student discipline mirrors the lower rates of discipline for younger students in the population, but created a difficulty in making statements regarding student grade as a contributor to discipline severity. Response rate challenges linked to sensitivity of data presented a major limitation to the present study usually through concerns of viable power.

Type II Error

When studies have low power, the chance of Type II (false negative) Error increases. The current study met the Kreft (1996) power recommendations for multi-level modeling with more than 30 groups with an average of more than 30 data points in each for overall and subjective discipline, but not for objective discipline. None of the discipline types approximated the more modern recommendation for multi-level modeling with a focus on cross-level interactions (effects of a level 2 variable on the dependent variable through a level 1 variable) of at least 100 groups with 10 data points in each (Maas & Hox, 2002). Hence, despite the lack of formal power calculations, the discipline severity analyses likely suffered from a high risk of Type II Error. Similarly,

the RDD calculations for power suggested a need for at least 20 additional participants for a large effect size and closer to 300 for a medium effect size.

With an undeniable lack of power in at least some of the current study, relationships and findings required a very high effect size for statistical analyses detection. Smaller, yet meaningful, effect sizes are typical of research involving human-subjects (Field, 2013). Considering the impact of changes in discipline severity by race/PRS and/or the RDD, even small effect sizes are critical to reveal. Administrators can work to lessen the effect of implicit bias on their decisions which creates a potential solution area if implicit bias is found responsible for some of the differences in discipline severity by race/PRS and/or RDD.

Study Design

The study design contained some challenges that became study limitations such as the vagueness of racial perceptions, the influence of explicit bias, the two- rather than three-level analyses, and the localization to PA. Although the researcher attempted to design the best possible study to investigate the research questions, these limitations remained impossible to remove. As such, readers should consider these limitations when accepting and developing practices based on the results and findings presented.

When collecting racial data from students, schools typically collect data reported by parents and limit such data to the following categories: Black/African American; Asian or Pacific Islander; Native American; Hispanic/Latino; White/Caucasian; and multiracial. Racial data is recorded when students enter the school system and schools rarely provide parents opportunities to update this data. As such, racial information often suffers from parental reporting bias or does not clearly reflect the race perceived by

109

others. In many studies and assessments using this data, researchers aim to evaluate the presence or absence of oppression by race—and typically perceived race. The recorded data provides unclear information regarding perceived race due to differences in racial perceptions of Hispanic/Latino students, Asian/Pacific Islander students, Native American students, and multiracial students. Of most concern to this study were multiracial students who very often identify as of Color but the designation does not provide enough information to rely on an assumption of such. Even inclusion of Hispanic/Latino students became a point of contention during the study due to schools' concerns over "White" Hispanic students. The study sought to account for this by decontextualizing race as race/PRS by adding an intermediate code for Hispanic students rather than simply White or of Color, but the schools' concern is valid. Nonetheless, even consideration of the three-level variable as PRS remains flawed as it does not necessarily reflect the students' phenome inclusive of race and traits. The schools' data regarding race was the only available data for use in this study, but did not provide an ideal measure.

Another concern was the lack of exploration of explicit versus implicit bias with regards to discipline differences by race/PRS. Although the researcher chose to assume most school administrators demonstrated only implicit biases due to the legal ramifications possible for those charged with explicit bias (in line with Nance, 2016), consideration of explicit bias in administrators is warranted. Administrators may explicitly believe in stereotypes and act on such beliefs with intention. For example, if a principal thought that Black students required more stringent discipline to succeed in K12 schools—even if she believed Black students could succeed as well or better than White

students—she might exhibit explicit biases which contribute to differences in discipline severity by race/PRS. Due to concerns related to the invasiveness and legal responsibility involved when measuring explicit bias, the researcher did not measure this kind of bias. Nonetheless, parsing implicit and explicit bias as contributors to racial discipline discrepancies remains a vital next step for developing a meaningful understanding of administrators' behaviors when making disciplinary decisions.

As discussed earlier, use of a two-level model using only students and administrators rather than a three-level model incorporating districts/schools greatly reduced the range and confidence of data interpretations. Without a school/district level, the models were incapable of understanding the effects of policy and contextual pressures on administrators and subsequently on students. Especially with regards to findings comparing subjective and objective discipline, potentially a school/district level could parse policy (seen in objective discipline) from other elements of interest at the school/district level. As a two-level model, such implications taken from the results remain questionable.

The restriction to only PA schools/districts built into the study design added a potential lack of generalizability. Although PA provided an excellent model of US schools (see Chapter 3; Sample), contextual concerns related to PA potentially distorted the findings when using this limited sample. Even if the results remained generalizable to US schools, international schools potentially suffer from much different racial or ethnic biases due to the history of the individual country; and so, non-US schools should not attempt to generalize data from this study.

Threat of Type I Error

Like most, this study also involves a risk of Type I, or false positive, error. Although alpha levels were set at a typical .05 level reducing the threat of Type I error to 5%, separation of the data into discipline types using the same administrator data potentially increased familywise error. As the datasets were not identical when conducting repeat analyses, the researcher could not definitively determine risk level for familywise error. Fortunately, most of the significant results were significant at an alpha level of less than .001, which decreases the Type I error risk to 0.1% for those findings. Using different datasets or reducing the overall alpha-level could lessen the risk of Type I error in this kind of study.

<div align="center">

Recommendations for Practice

</div>

Administrators' implicit bias presents a noteworthy concern for decision-based school discipline. Schools should consider creating less subjective discipline systems to reduce the effects of implicit bias (Smolkowski, Girvin, McIntosh, Nese, & Horner, 2016). Schools could develop a blinded discipline appropriation system where student information is removed from decision-makers' awareness when possible. Use of an uninvolved discipline decision-maker from a different school or grade could facilitate a blinded system. At the teacher level, schools might employ structured decision-making protocols already shown to decrease ODRs and discipline disparities by race/PRS (Girvin, Gion, McIntosh, & Smolkowski, 2016; Yusuf, Irvin, & Bell, 2016). These systems typically provide a "road map" for discipline that decreases the potential level of subjectivity in decision making. Use of guidelines of acceptable discipline for different infractions based on the instance of infractions would assist in the development of a

disciplinary decision structure (Mukuria, 2002). Although some decision-making remains warranted and even ideal with regards to discipline, removing emotionality and developing guidelines will remove the high-cognitive load conditions that kindle unhindered implicit bias in decision-making.

Policy-makers should develop potentially-objective discipline decisions into policy-driven disciplinary actions. This has already begun in some districts through state and federal mandates (Losen, & Haynes, 2016). Some policies reduce exclusionary discipline through non-allowance of exclusionary discipline for more minor infractions. In the recently released Every Student Succeeds Act (ESSA), specific requirements exist related to lowering overall and racially inequitable exclusionary discipline (The Leadership Conference Education Fund, 2016). This act bridges research with policy by following the three suggestions for discrepancy reduction recommended by Losen and Haynes (2016): (1) collect, analyze, and report school data at least annually, (2) implement alternatives to exclusionary discipline, and (3) align discipline with academics. ESSA calls for annual reporting of school discipline data disaggregated by race, which has repeatedly been called for and found effective (Losen & Haynes, 2015; Skiba, Horner, Choong-Geun, Rausch, May, & Tobin, 2011; Skrla, Scheurich, Garcia, & Nolly, 2004). ESSA requires plan development for the reduction of exclusionary discipline overuse and, "use of aversive behavioral interventions that compromise student health and safety" ESSA, Section 1111(g)(1)(C)(i-iii). Finally, ESSA calls for a discipline-academics alignment by creating school-wide positive behavior interventions and supports, bullying and harassment prevention, school-dropout and reentry programming, and other preventative strategies to keep kids in the classroom. Such

113

increased measures of accountability and concern related to exclusionary discipline pose valid methods of implicit bias mitigation (Cate, Krolak-Schwerdt, & Glock, 2015). Schools should take initiative to perform such data analyses prior to policy-driven requirements in order to facilitate effective data collection, analysis, and reporting procedures that make sense to all those involved. This process should at least begin with shared decision-making where data procedures are discussed with all involved personnel rather than only administrators to help develop a system personnel want to follow and understand (Nishioka, Shigeoka, & Lolich, 2017).

Schools should consider adopting culturally relevant positive behavior support systems that incorporate culturally responsive and relevant teaching, curriculum, and discipline into the school's culture (Vincent, Randall, Cartledge, Tobin, & Swain-Bradley, 2011). Many schools and districts already use positive behavior support systems, but are unaware of their inability to create racially equitable school discipline (Tobin, Sugai, & Colvin, 1996). Although this alone is unlikely to remove the inequities in exclusionary discipline by race, such systems might begin to develop a culture that provides personnel with a better understanding of the students served and lessen bias through an added information methodology (see Mann & Ferguson, 2015). Furthermore, discipline support systems in general often advocate for increased decision-making time when reporting ODRs and when deciding on the appropriate behavior consequences. Although often intended as a method of removing emotionality or providing time for restorative practice conferences, the increased decision-making time also allows for cognitive bias mediation.

Schools can implement bias intervention systems similar to those discussed in Kahn, Goff, and Glaser (2016) and Devine, Forscher, Austin, and Cox (2012). At a minimum, schools should expose professionals to implicit bias and explain the construct and how it could act on behaviors (Capers, Clinchot, McDougle, & Greenwald, 2017; Staats, 2015). To maintain trust and calm, exposure should be followed-up with bias mediation and mitigations strategies (see Chapter 2). Together, awareness/exposure and strategy education hold the potential to develop the knowledge necessary for school professionals to recognize bias (both implicit and explicit) and begin to develop a self-awareness of the effects of implicit bias on their practices. Researchers know implicit bias affects racial disparities in exclusionary discipline, and must continue to research methods of removing and/or mitigating bias in teachers and other school professionals (Warikoo, Sinclair, Fei, & Jacoby-Senghor, 2016). Change needs to occur in order to protect the equity of school justice.

Recommendations for Research

Based on the findings and limitations of this study, future research suggestions fit into three categories: (1) design enhancement, (2) measurement, and (3) complementary studies. Research enhancing the design of the present study should focus on including a third level of HLM to include school/district level factors. To add this level effectively, the study should include a larger sample stemming from more US states. Other studies should consider using different measures and analysis methods for race/PRS variables. Use of dummy coding to compare Hispanic to White and Black to White populations separately would produce more distinct results regarding relationships between IAT and discipline as experienced by each unique population. Furthermore, researchers must

consider using dual measures for race and PRS with race defined by participant self-identification and PRS defined by perceptions of the participant. The study design could benefit from a more specific population to produce a better response rate as well. If a single district design was employed in a district concerned with discipline equity, a higher response rate could be expected. Moreover, researchers should consider ways to increase response rates in larger populations such as better incentives, easier methods of participation, or taking administrator measurements prior to the public release of discipline records.

Measurement studies should investigate national discipline figures separated by objective and subjective discipline types to determine whether a larger than expected discipline gap appears in subjective disciple. Other studies of discipline decisions should also incorporate separate measures for each discipline type while accounting for issues related to familywise error. Additional measurement studies could evaluate the validity and/reliability of RDD at the administrator level as a localized measure of discipline discrepancy. Studies should compare RDD to school-, district-, state-, and country-wide racial discipline gaps to better understand the relationship between discrepancy rates at each level.

Complementary studies should begin by considering a replication of this study inclusive of the previously described design enchantments with delineation between implicit and explicit bias. Anonymity is critical in a study of this nature due to legal implications, but use of more popular explicit bias measures only accounts for symbolic explicit bias which may not result in the same level of legal implications. The results of the current study indicated the potential of interventions to remediate implicit bias at the

116

administrator level to diminish the inequitable race/PRS to discipline severity relationship; hence, a quasi-experimental repeated-measures study providing implicit bias mitigation training for administrators comparing the relationship between implicit bias and the race/PRS-discipline severity relationship before and after the training remains necessary. Both Devine and Kahn offer similar studies outside of the school context or with no measure of school discipline effects which provide exemplary designs for such a study (Devine, Forscher, Austin, & Cox, 2013; Kahn, Goff, & Glaser, 2016). Additionally, researchers should expand considerations of the impact of implicit biases in schools to include other relationships within the chain of discipline including between the referring teacher and the decision-making principal, between the assistant principal and head principal, between the principal and superintendent, and between any school professional and a student's family. Furthermore, studies might begin to explore the level of discipline policy adherence (strict vs. loose) employed by different schools, districts, and/or principals in order to evaluate the true objectivity of infractions categorized as such. With the addition of the suggested research, this study could facilitate meaningful changes with the potential to vastly improve equity in school discipline throughout the US.

Summary

The current study presented administrator implicit bias as a possible mechanism for the school racial discipline gap and evaluated this claim through student discipline severity and RDD as they related to administrator implicit bias. The findings suggested that implicit bias acted on the race/PRS to discipline severity relationship only when discipline decisions were subjective in nature. Schools need to work towards disciplinary

117

equity and may do so with interventions related to administrators' implicit bias or the level of decision-making involved in discipline. The researcher hopes that this study will lead to more studies of implicit bias and school discipline which consider the school administrator as a potential source of bias to allow for mediation of implicit bias-based discipline discrepancies beyond the teacher level.

TABLES

Table 1

Table of Abbreviations

Abbreviation	Definition
ADEXP	Administrator Experience (in years)
ADRACE	Administrator Race
CRDC/OCR	Civil Rights Data Collection/ Office of Civil Rights
ESSA	Every Student Succeeds Act
FRL	Free or Reduced Lunch
FUM	Fully Unconditional Model
HLM	Hierarchical Linear Modeling
IAT	Implicit Associations Test
ICC	Intra-class Correlation
INFLVL	Infraction Level
ISS	In-School Suspension
ODR	Office Disciplinary Referral
OLS	Ordinary Least Squares
OSS	Out of School Suspension
PRS	Phenotypic Racial Stereotypicality
RDD	Racial Discipline Discrepancy
RQ	Research Question
SES	Socio-Economic Status

Table 2

Coding of Infraction Level and Type Using All Possible Codes in the PA Safe Schools Database

Inf. Level	PA Safe Schools Infraction Description		
1	Reckless Endangering (S)	Criminal Trespass (S)	Disorderly Conduct (S)
2	Robbery (O)	Theft (S)	Burglary (O)
	Vandalism (S)	Failure of Disorderly Persons to Disperse upon Official Order (S)	
3	Indecent Exposure (S)	Open Lewdness (S)	Sexual Harassment (O)
	Racial/Ethnic Intimidation (S)	Minor Altercation (S)	Stalking (O)
	Threatening Student (O)	Bullying (O/S)	TOWARDS STUDENTS
	Obscene and other sexual materials and performances (S)	All Other Forms of Harassment (O/S)/Intimidation (S)	
4	Indecent Exposure (S)	Open Lewdness (S)	Sexual Harassment (O)
	Racial/Ethnic Intimidation (S)	Minor Altercation (S)	Stalking (O)
	Threatening School Official (O)	Bullying (O/S)	TOWARDS STAFF
	Obscene and other sexual materials and performances (S)	All Other Forms of Harassment (O/S)/Intimidation (S)	
5	(Statutory) Sexual Assault (O)	Rape (O)	Fighting (O)
	(Aggravated) Indecent Assault (O)	(Simple/Aggravated) Assault (O)	Unlawful Restraint (O)
	Involuntary Sexual Deviate Intercourse (O)		**TOWARDS STUDENTS**
6	Sexual Assault (O)	Rape (O)	Fighting (O)
	(Aggravated) Indecent Assault (O)	(Simple/Aggravated) Assault (O)	Unlawful Restraint (O)
	Involuntary Sexual Deviate Intercourse (O)		**TOWARDS STAFF**
7	Sale, Possession, Use, or Under the Influence of Alcohol (O)	Sale, Possession, Use, of Tobacco (O)	
8	Possession/Use of a Controlled Substance (O)	Sale/Distribution of a Controlled Substance (O)	
9	Attempt/Commit Murder/Manslaughter (O)	Arson (O) Bomb Threats (O)	
	Terroristic Threats (excluding bomb threats) (O)	Possession of Handgun or Rifle/Shotgun (O)	
	Possession of BB/Pellet Gun (O)	Possession of Other Firearm (O) Possession of Other Weapon (O)	
	Possession of Knife or Cutting Instrument (O)	Possession of Explosive (bomb, missile, etc.) (O)	

O = Objective; S = Subjective

120

Table 3

Comparison of 2011-2012 PA and National Exclusionary Discipline for Students of Color

Discipline	PA	National
One or more in-school suspensions	38%	53%
Only one out-of-school suspension	52%	57%
More than one out-of-school suspension	60%	63%
One or more out-of-school suspensions	56%	60%
Expulsions with educational services	44%	57%
Expulsions without educational services	67%	56%
Expulsions with or without educational services	53%	57%
Expulsions under zero-tolerance policies	47%	53%
Referral to law enforcement	41%	51%
School-related arrests	48%	54%

Note. Percentages represent students of Color receiving discipline over all students receiving discipline specified.

Table 4

Schematic depiction of the Implicit Associations Test.

Sequence	1	2	3	4	5
Task Description	Images	Words	Combined 1	Image Switch	Combined 2
Category Instructions	Black White	Pleasant Unpleasant	Black Pleasant White Unpleasant	White Black	White Pleasant Black Unpleasant
	Black A	Lovely	Black C	Black B	White A
	Black B	Frown	Pleasant	White D	Pleasant
	White A	Pleasant	Frown	White A	Frown
Sample	Black C	Happy	White A	Black C	Lovely
Stimuli	White B	Depressed	Depressed	White B	Black B
	Black D	Filthy	Smile	Black A	White A
	White C	Unpleasant	Filthy	Black D	Filthy
	White D	Smile	Black D	White C	White C

Note. Letters denote different images in each defined category.

Table 5

IAT Word and Image Stimuli

	Category A	Category B
Words	Positive	Negative
	Cheer	Agony
	Freedom	Crash
	Gentle	Death
	Happy	Disaster
	Health	Evil
	Honest	Filth
	Honor	Grief
	Laughter	Hatred
	Loyal	Rotten
	Lucky	Stink
	Peace	Tragedy
	Sunrise	Ugly
Images	White	of Color

Table 6

Comparison of Participating and Declining Districts

	Declining Districts				Participating Districts				Comparison	
	$\bar{x}$	S.D.	S.E.	Med.	$\bar{x}$	S.D.	S.E.	Med.	t (136)	Sig.
Student Enrollment	5174	11896.3	1039.4	3407	8389	3950.0	1492.9	6512	-0.71	.48
Professional Personnel	407	820.2	71.7	279	685	139.0	120.6	531	-0.90	.37
Student Demographics										
% White	63	20.4	1.8	70	61	25.2	9.5	66	0.23	.81
% Black	16	17.7	1.5	9	9	5.1	1.9	8	2.91*	.01
% Hispanic	12	11.9	1.0	9	18	22.3	8.4	8	-0.69*	.52
% of Color	29	19.7	1.7	21	27	25.7	9.7	4	0.18	.86
Student Discipline										
Incidents	202	576.2	50.3	63	218	275.5	104.1	60	-0.08	.94
% Offenders	3	3.3	0.3	2	2	1.7	0.6	2	1.06	.29
Administrators										
Count	18	27.3	2.4	13	31	12.9	4.9	24	-1.19	.24
Salary (in thousands)	104	14.7	1.3	101	111	14.0	5.3	107	-1.29	.20
Years of Experience	18	4.2	0.4	18	18	1.4	0.5	17	0.11*	.92
Years in District	11	3.7	0.3	11	12	2.0	0.7	13	-0.85	.40
Education Level	5	0.2	< 0.1	5	5	0.2	< 0.1	5	0.56	.58

Note: Variables in italics denote those with both significant differences found via *t*-test and non-equal medians.

*Levene's Test for Equality of Variances was significant at the 0.05 level. Equal variances not assumed and compensating *t* tests were run with 15.9, 6.2, and 12.6 d.f., respectively.

Table 7

Student Level Descriptive Statistics with Data Comparisons between Participating and Declining by Decision Type

	Declining Administrator					Participating Administrator					Comparison	
	n	$\bar{x}$	S.D.	S.E.	Med.	n	$\bar{x}$	S.D.	S.E.	Med.	t	Sig.
All Discipline												
FRL Eligibility	2468	1.54	0.83	.02	2	3661	1.29	0.94	.02	2	11.00	<.001
Race/PRS	2362	0.96	0.69	.01	1	3432	0.79	0.80	.01	1	8.58	<.001
Grade	2468	8.65	2.70	.05	9	3663	9.58	1.95	.03	10	-14.68	<.001
Infraction Level	2468	1.60	1.47	.03	1	3663	1.74	1.48	.03	1	-3.81	<.001
Discipline Severity	2468	3.00	1.36	.03	2	3662	2.94	1.40	.02	2	1.84	.066
Exclusionary Discipline	2468	0.36	0.48	.01	0	3663	0.43	0.50	.01	0	-5.52	<.001
Subjective Discipline												
FRL Eligibility	2352	1.54	0.84	.02	2	3496	1.28	0.95	.02	2	10.85	<.001
Race/PRS	2251	0.96	0.69	.02	1	3278	0.78	0.80	.01	1	8.58	<.001
Grade	2352	8.70	2.68	.06	9	3498	9.59	1.95	.03	10	-13.97	<.001
Infraction Level	2352	1.20	0.59	.01	1	3498	1.40	0.88	.02	1	-10.13	<.001
Discipline Severity	2352	2.83	1.29	.03	2	3498	2.79	1.32	.02	2	1.29	.198
Exclusionary Discipline	2352	0.31	0.46	.01	0	3498	0.39	0.49	.01	0	-6.68	<.001
Objective Discipline												
FRL Eligibility	248	1.65	0.73	.05	2	401	1.56	0.82	.04	2	1.53	.126
Race/PRS	237	0.98	0.71	.05	1	373	0.94	0.79	.04	1	0.66	.510
Grade	248	8.38	2.64	.17	9	401	9.38	1.84	.09	9	-5.28	<.001
Infraction Level	248	5.64	1.58	.10	5	401	5.46	1.48	.07	5	1.43	.153
Discipline Severity	248	4.81	0.66	.04	5	400	4.59	1.10	.06	5	3.13	.002
Exclusionary Discipline	248	0.95	0.22	.01	1	401	0.89	0.32	.01	1	3.16	.002

Note: Variables in italics denote those with both significant differences found via *t*-test and non-equal medians.

S.E. = standard error of the mean; Med. = median

Table 8

Enrollment demographics for schools and districts represented in the participating sample.

	Mean	SD	Median	Min	Max
Districts					
Enrollment	7,673	4,180.3	6,045	2,661	13,539
% Black	13	11.6	9	3	39
% Hispanic	17	20.8	8	3	60
% White	58	25.5	65	13	86
% Economically Disadvantaged	43	24.5	39	14	91
% English Language Learners	4	5.2	4	<1	16
% Special Education	15	2.0	16	11	17
Schools with Disciplinary Data					
Enrollment	867	674.5	558	296	2,987
% Black	12	5.9	11	3	23
% Hispanic	32	30.2	13	2	81
% White	41	29.9	46	2	89
% Economically Disadvantaged	58	32.8	52	13	98
% English Language Learners	9	7.5	10	<1	23
% Special Education	17	3.2	17	11	24
Schools with Participating Administrators					
Enrollment	1,002	781.2	649	296	2,987
% Black	10	4.7	10	3	21
% Hispanic	29	26.2	22	2	77
% White	49	25.9	61	4	89
% Economically Disadvantaged	51	28.9	42	13	95
% English Language Learners	6	6.1	4	<1	23
% Special Education	16	3.7	17	9	26
Schools with Participating Administrators in Schools with Disciplinary Data					
Enrollment	1,155	802.1	795	296	2,987
% Black	10	6.1	8	3	21
% Hispanic	19	26.2	5	2	72
% White	55	28.0	64	4	89
% Economically Disadvantaged	42	30.8	29	13	95
% English Language Learners	5	7.2	2	<1	23
% Special Education	16	2.2	17	11	19

Table 9

Administrator Level Descriptive Statistics

Variable	Med.	Mean	S.E.	SD	Min.	Max.
Race	0.00	0.08	0.04	0.26	0.00	1.00
Experience	4.00	5.64	0.79	4.80	1.00	18.00
IAT Score	-0.37	-0.37	0.06	0.35	-1.16	0.26
ED	38.00	83.97	39.35	180.95	1.00	1120.00
RDD	15.52	31.53	5.00	27.38	6.88	91.01

Note: For all variables $n = 39$, except RDD and subcategories where $n = 30$ due to missing data or less than 10 instances of exclusionary discipline.

ED = Exclusionary Discipline

Table 10a

Student Level Variable (and Variable Component) Correlation Matrix

	2.	3.	4.	5.	6.
All Discipline					
1. FRL Eligibility	.36***	-.20***	.00	.25***	.24***
2. Race/PRS	-	-.12***	-.05***	.18***	.16***
3. Grade	-	-	-.04**	-.15***	-.18***
4. Infraction Level	-	-	-	.39***	.34***
5. Discipline Severity	-	-	-	-	*.92***
6. ED	-	-	-	-	-
Subjective Discipline					
1. FRL Eligibility	.35***	-.20***	-.09***	.25***	.232**
2. Race/PRS	-	-.12***	-.15***	.18***	.15***
3. Grade	-	-	-.02	-.15***	-.18***
4. Infraction Level	-	-	-	.11***	.13***
5. Discipline Severity	-	-	-	-	*.91***
6. ED	-	-	-	-	-
Objective Discipline					
1. FRL Eligibility	.36***	-.14**	-.14***	.00	.01
2. Race/PRS	-	-.08*	-.13**	.09*	.09*
3. Grade	-	-	.07	-.02	-.07
4. Infraction Level	-	-	-	.20***	.12**
5. Discipline Severity	-	-	-	-	*.90***
6. ED	-	-	-	-	-

Note: Values represent Pearson's *r* with 2-tailed significance testing. Italics denote significant correlations above the .60 criterion for moderate and above correlations.
ED = Exclusionary Discipline
* *p* < .05, ** *p* < .01, *** *p* < .001

Table 10b

Administrator Level Variable (and Variable Component) Correlation Matrix

	2.	3.	4.	5.
1. Experience	-.06	-.20	.10	-.05
2. Race	-	.08	*.61****	.29
3. IAT Score	-	-	.07	-.09
4. ED- of Color	-	-	-	.38*
5. RDD- of Color	-	-	-	-

Note: Values represent Pearson's *r* with 2-tailed significance testing. Italics denote significant correlations above the .60 criterion for moderate and above correlations. Reduced font size denotes anticipated correlations.

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

Table 11a

Models of Discipline Severity for All Discipline Types

Effects	Coefficient	SE	Variance Component	Reliability Estimate
Fully Unconditional Model				
Intercept	3.02	0.14	0.66	0.89
Level 1 σ^2 remaining	-	-	1.49	-
Hypothesized Model				
Intercept	-	-	0.56	0.81
Intercept	2.91	0.13	-	-
Race/PRS slope	-	-	< 0.00	0.04
Intercept**	0.09	0.02	-	-
IAT*	-0.11	0.04	-	-
Level 1 σ^2 remaining	-	-	1.50	-
Final Model				
Intercept	-	-	0.52	0.84
Intercept	2.79	0.13	-	-
Race/PRS slope	-	-	< 0.00	0.11
Intercept	0.05	0.03	-	-
IAT	-0.10	0.05	-	-
FRL slope***	0.13	0.02	F	F
Inf. Level Slope***	0.41	0.03	F	F
Level 1 σ^2 remaining	-	-	1.15	-

Note. Mean intercept significance is not shown to highlight only effects meaningfully interpretably significance.

* $p < .05$, ** $p < .01$, *** $p < .001$, F = Fixed effect

Table 11b

Models of Discipline Severity for Subjective Discipline

Effects	Coefficient	SE	Variance Component	Reliability Estimate
Fully Unconditional Model				
Intercept	2.88	0.13	0.64	0.89
Level 1 σ^2 remaining	-	-	1.28	-
Hypothesized Model				
Intercept	-	-	0.56	0.83
Intercept	2.80	0.13	-	-
Race/PRS slope	-	-	< 0.00	0.06
Intercept**	0.08	0.02	-	-
IAT*	-0.09	0.04	-	-
Level 1 σ^2 remaining	-	-	1.29	-
Final Model				
Intercept	-	-	0.49	0.82
Intercept	2.65	0.12	-	-
Race/PRS slope	-	-	0.01	0.16
Intercept*	0.06	0.03	-	-
IAT*	-0.08	0.04	-	-
FRL slope***	0.13	0.02	F	F
Inf. Level Slope***	0.42	0.06	F	F
Level 1 σ^2 remaining	-	-	1.17	-

Note. Mean intercept significance is not shown to highlight only effects meaningfully interpretably significance.

* $p < .05$, ** $p < .01$, *** $p < .001$, F = Fixed effect

131

Table 11c

Models of Discipline Severity for Objective Discipline

Effects	Coefficient	SE	Variance Component	Reliability Estimate
Fully Unconditional Model				
Intercept	4.48	0.14	0.31	0.48
Level 1 σ^2 remaining	-	-	1.12	-
Hypothesized Model				
Intercept	-	-	0.46	0.57
Intercept	4.26	0.18	-	-
Race/PRS slope	-	-	0.02	0.09
Intercept**	0.35	0.11	-	-
IAT	0.36	0.20	-	-
Level 1 σ^2 remaining	-	-	1.11	-
Final Model				
Intercept	-	-	0.53	0.61
Intercept	4.28	0.18	-	-
Race/PRS slope	-	-	0.01	0.05
Intercept*	0.19	0.09	-	-
Inf. Level Slope***	0.27	0.08	F	F
Level 1 σ^2 remaining	-	-	1.00	-

Note. Mean intercept significance is not shown in order to highlight only effects meaningfully interpretably significance.

* $p < .05$, ** $p < .01$, *** $p < .001$, F = Fixed effect

Table 12

Variance Statistics for Models of Discipline Severity

Statistic	FUM	Hypothesized Model	Final Model
All Discipline ICC = 0.31 HM = 15.57			
σ^2	1.49	1.50	1.15
τ_{00}	0.66	0.56	0.52
τ_{11}	-	< 0.01	< 0.01
Level 1 R^2	-	0.04	0.22
Level 2 R^2	-	0.14	0.21
Student Race/PRS Slope	-	0.87	NS
Subjective Discipline ICC = 0.33 HM = 16.45			
σ^2	1.28	1.29	1.17
τ_{00}	0.63	0.57	0.49
τ_{11}	-	< 0.01	0.01
Level 1 R^2	-	0.03	0.13
Level 2 R^2	-	0.09	0.21
Student Race/PRS Slope	-	0.89	0.89
Objective Discipline ICC = 0.22 HM = 7.02			
σ^2	1.12	1.11	1.00
τ_{00}	0.31	0.46	0.53
τ_{11}	-	0.02	0.01
Level 1 R^2	-	0.01[a]	0.11[a]
Level 2 R^2	-	NV	0.71[a]
Student Race/PRS Slope	-	NS	NS

Note: Pseudo R^2 values were only calculated as appropriate for the respective models. Values were calculated using the Snijders & Bosker (1994, 2012) calculation unless noted otherwise.

[a] Invalid result, recalculated using the Kreft & de Leeuw (1998) and Singer (1998) calculation.

ICC = Intra-class correlation; HM = Harmonic mean for n; NS = Effect not significant; NV = No valid result found

FIGURES
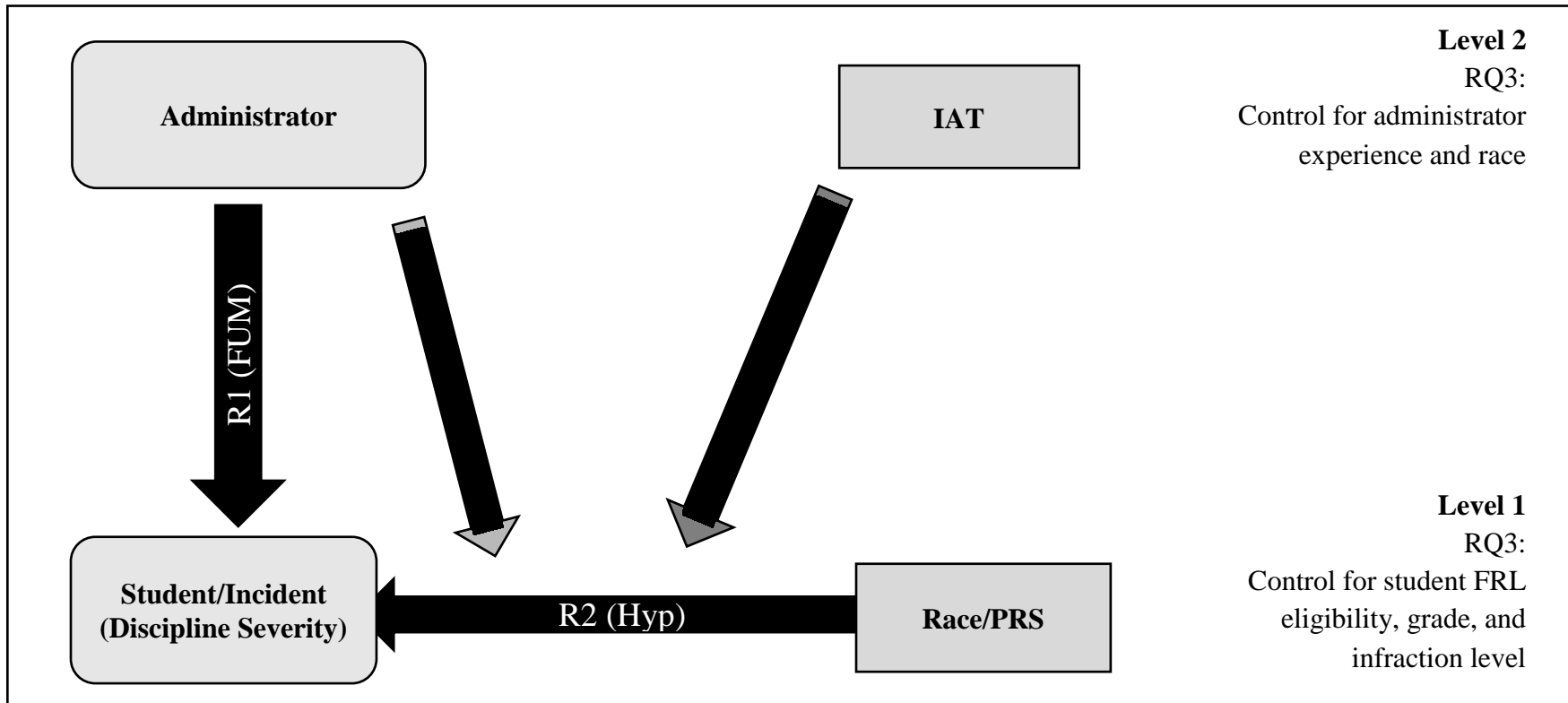
Figure 1. Schematic depiction of discipline severity research questions. Black arrows represent effects on means/intercepts and grey arrows represent effects on slopes.

134

Phase I

Invited school/districts to participated and obtained informed consent and site permission.

IRB Submission/ Approval

Phase II

Screened particpant datasets for complete information.

Surveyed Administrators.

Phase III

Collected datasets after 100 days of school completed.
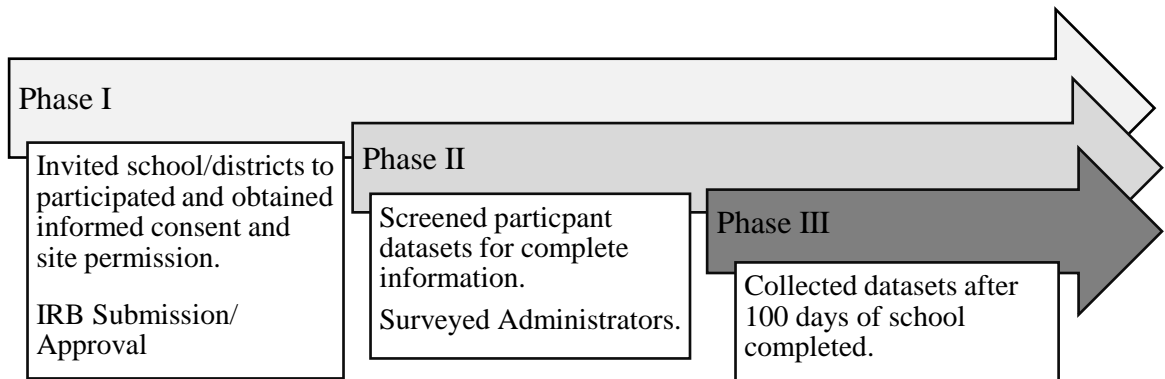
Figure 2. Phases of data collection.

135
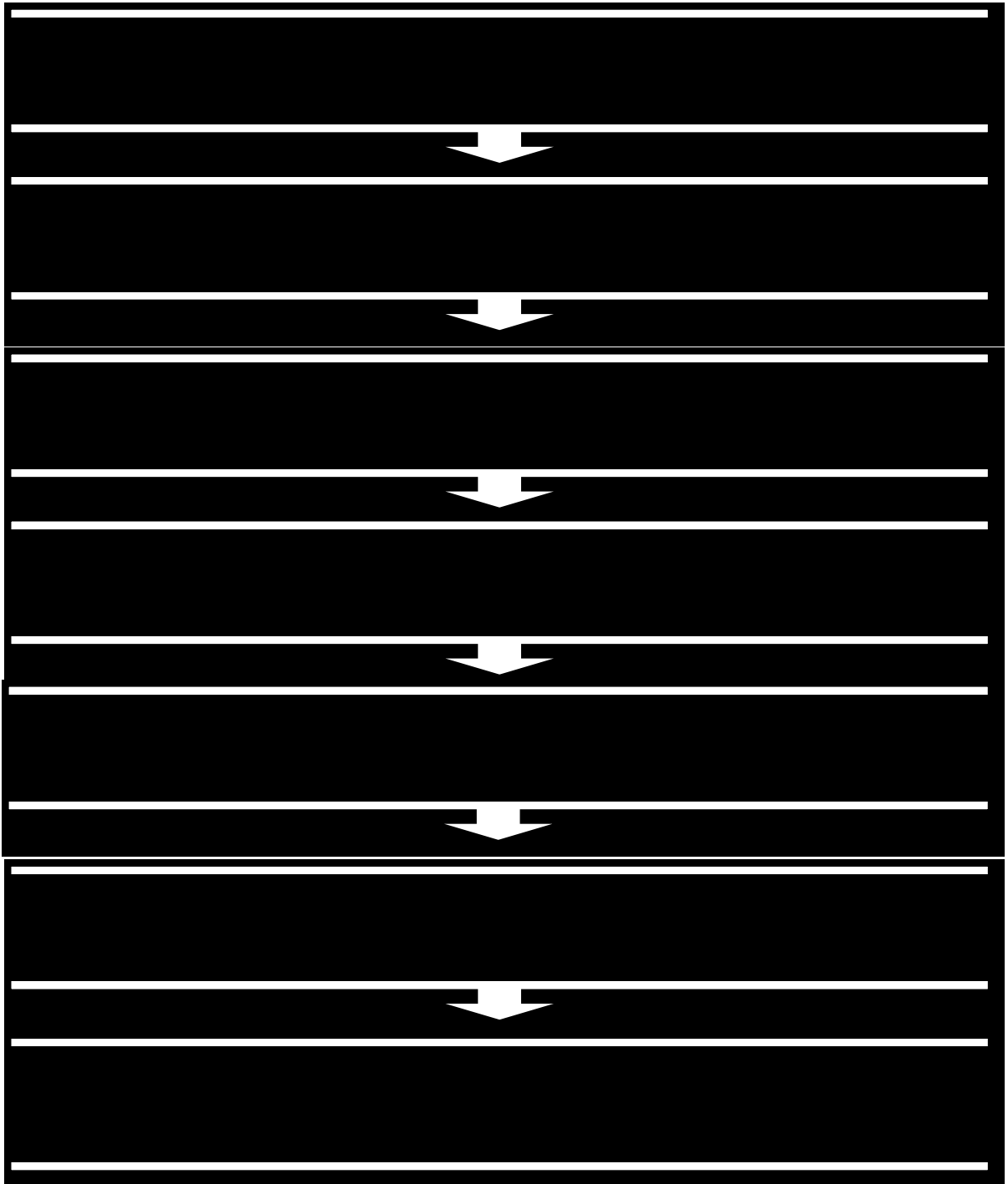
Figure 3. Modern IAT Race Sequence
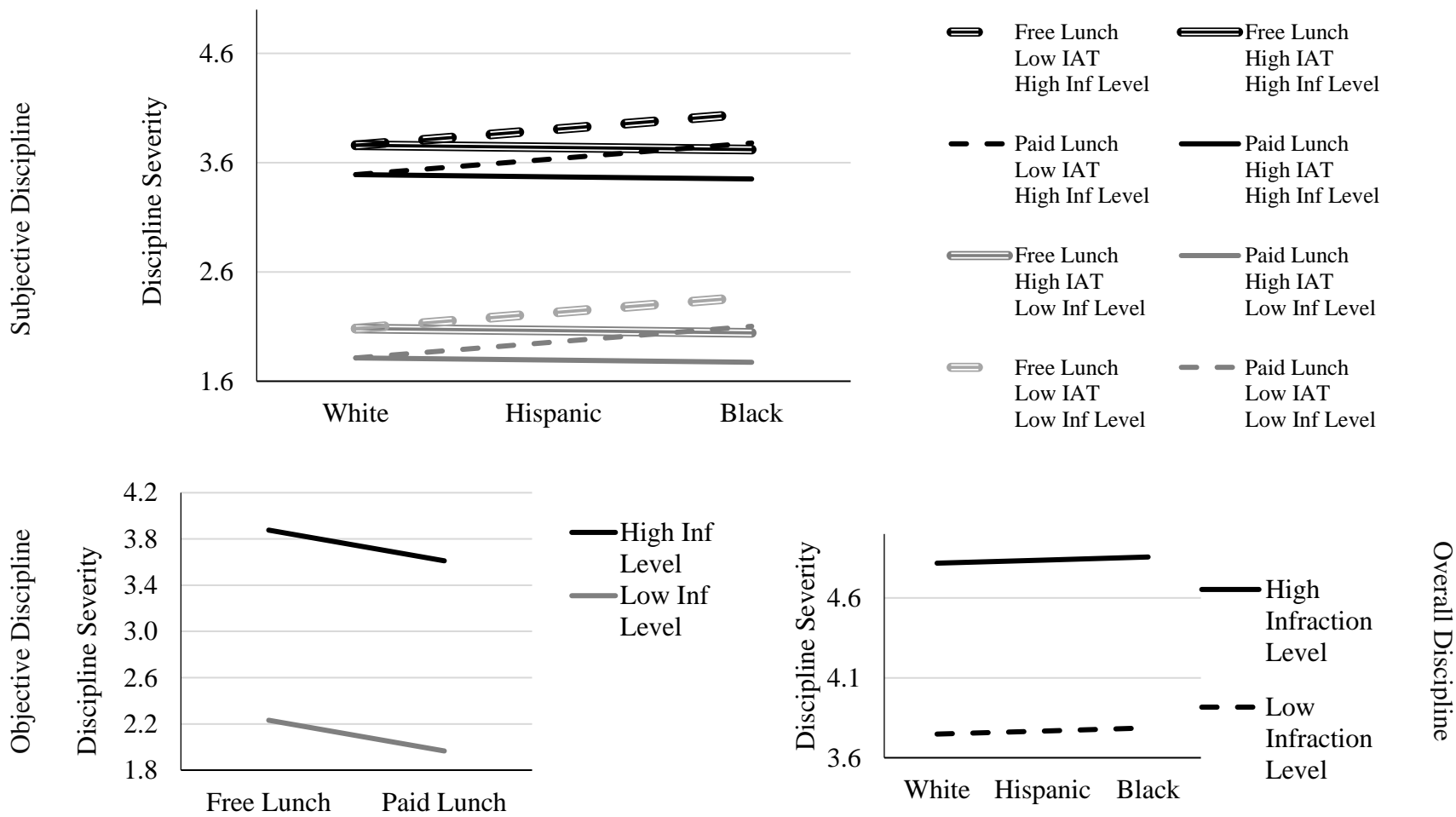
Figure 4. Graphical representations of final models for each discipline type category.

137
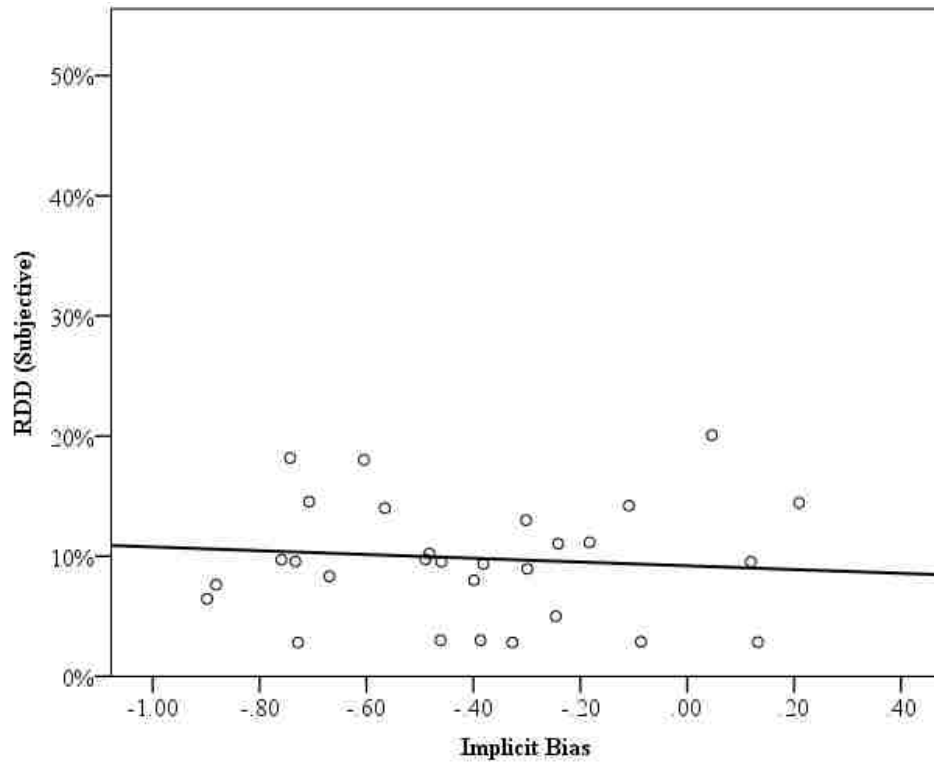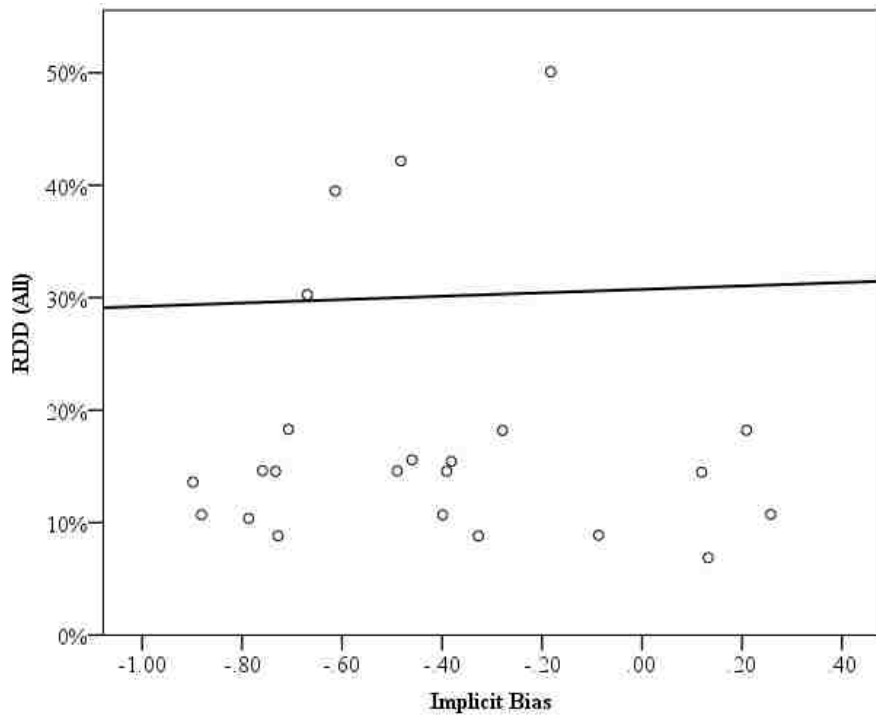
Figure 5. Racial Discipline Discrepancy (RDD) by IAT score separated into all discipline and subjective only discipline.
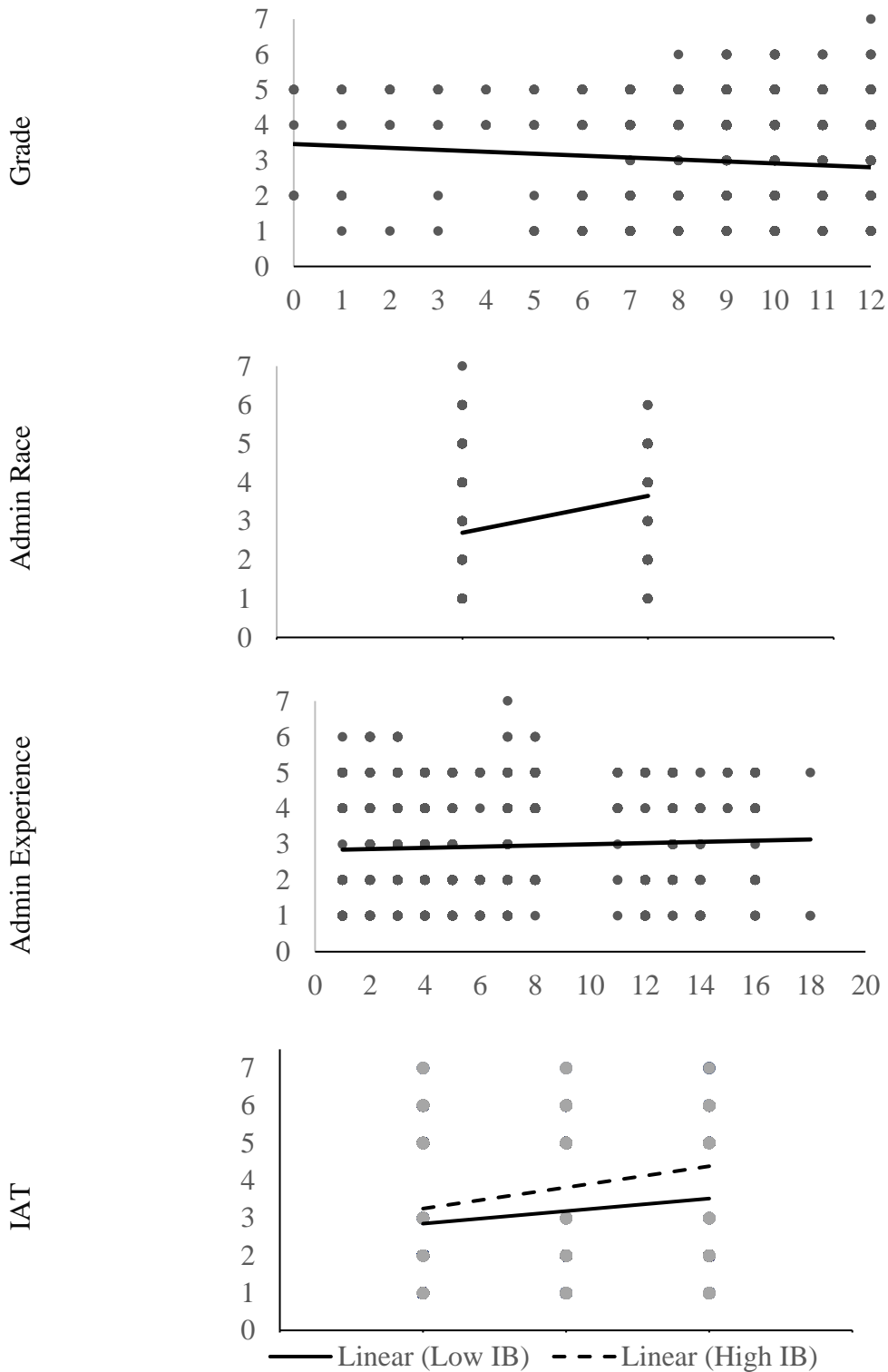
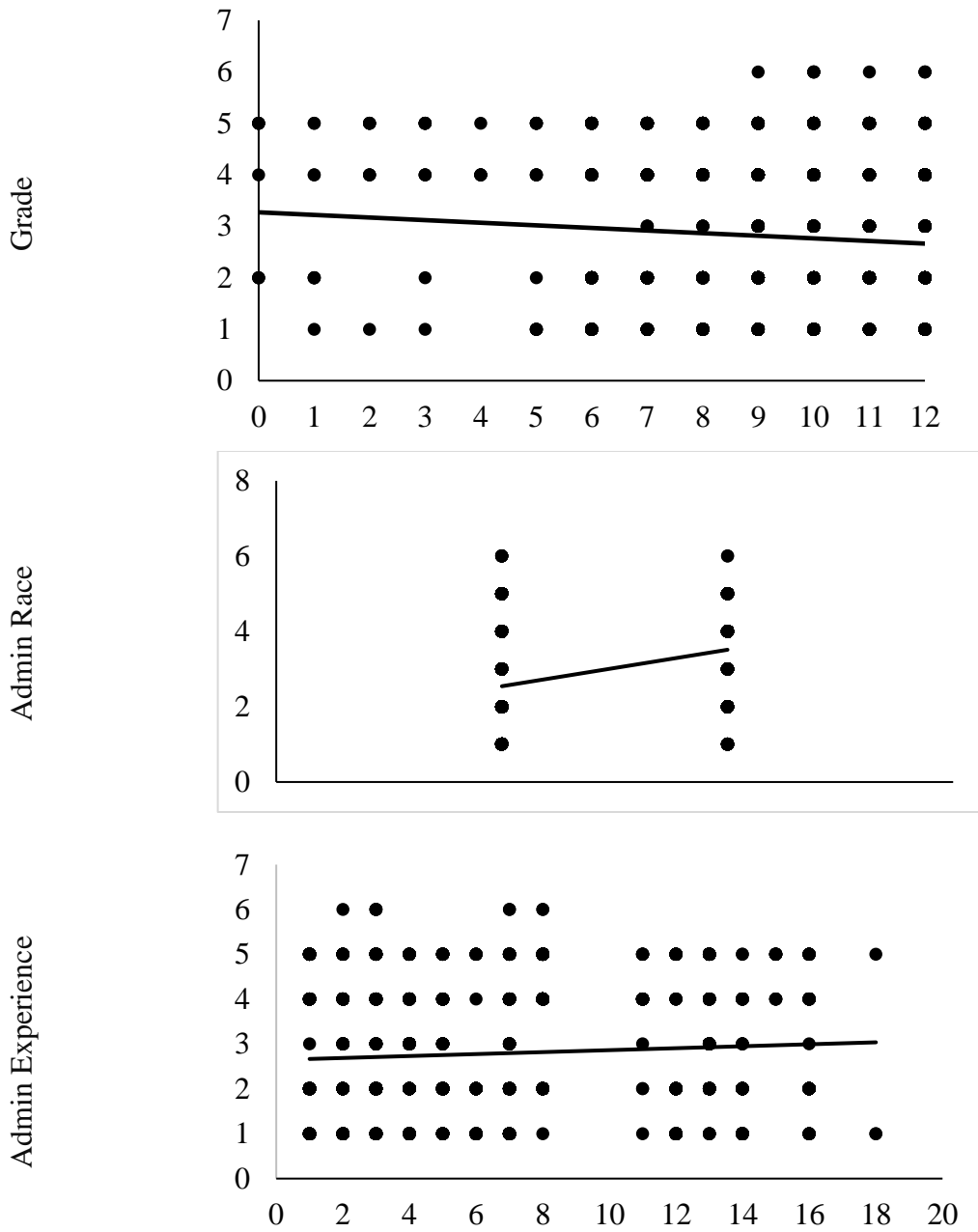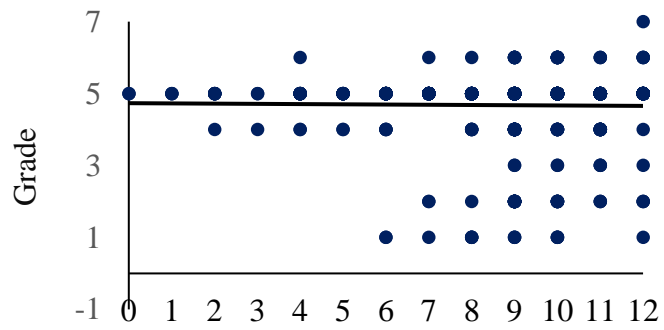*Figure 6.* Linearity assumption testing for non-significant effects in overall discipline severity model**.**
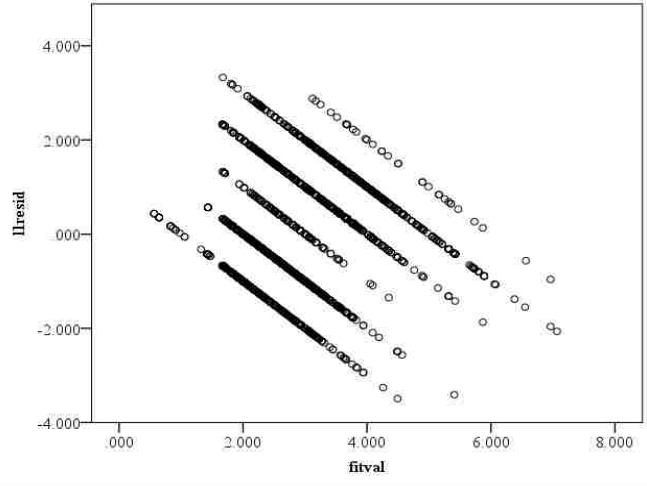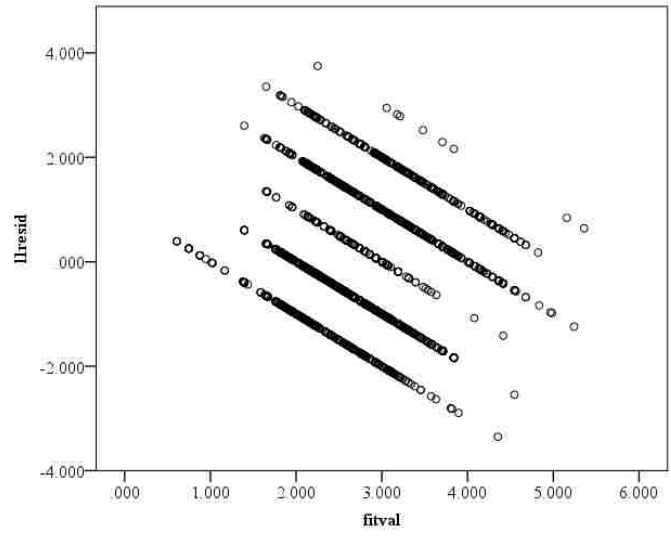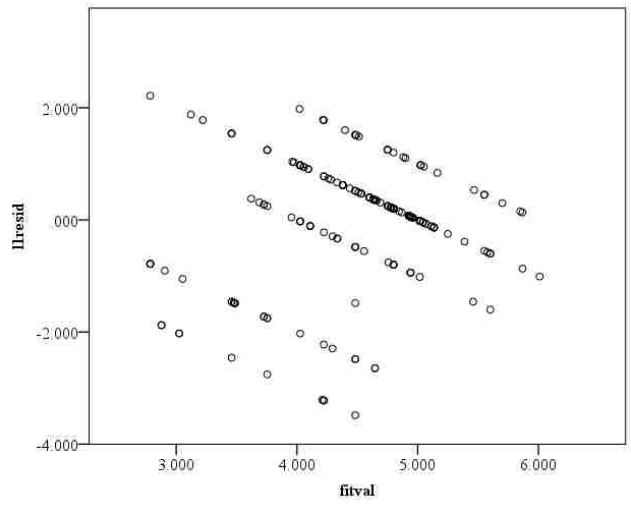
Figure 7. Graphical tests of level one residual homoscedasticity for discipline severity models.

Figure 8. Graphical tests of level two intercept residual homoscedasticity for discipline severity models.

All Discipline
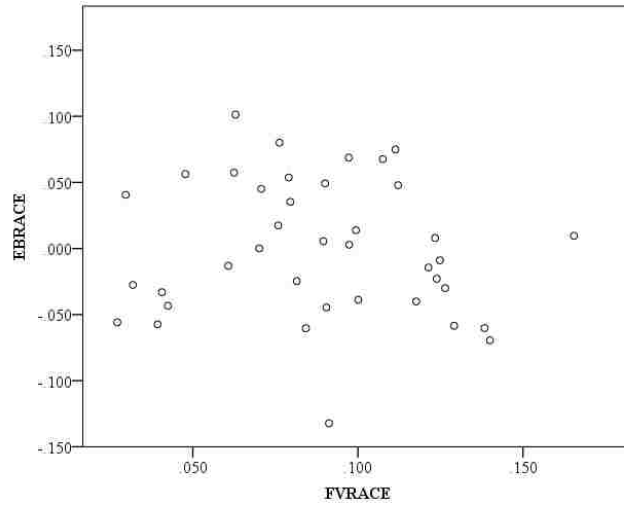
Subjective Discipline
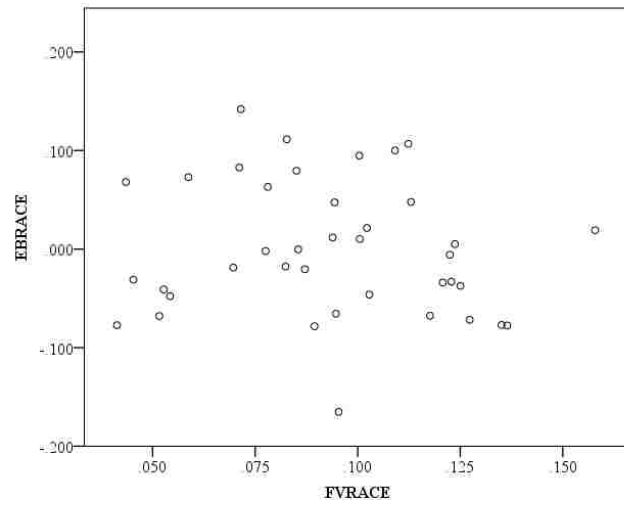
Objective Discipline
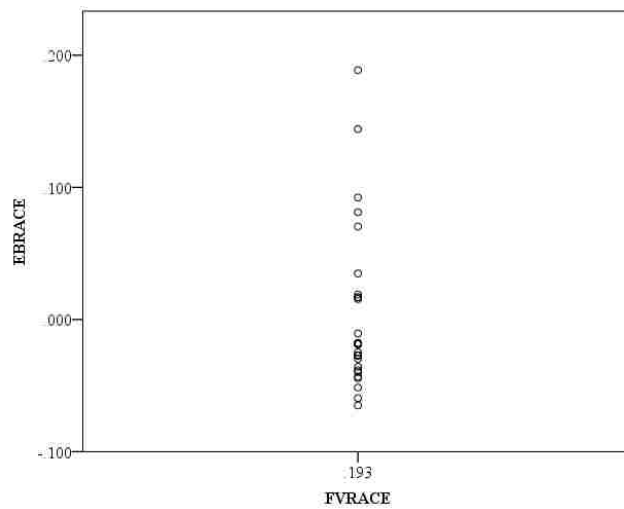
Figure 9. Graphical tests of level two race/PRS slope residual homoscedasticity for discipline severity models**.**
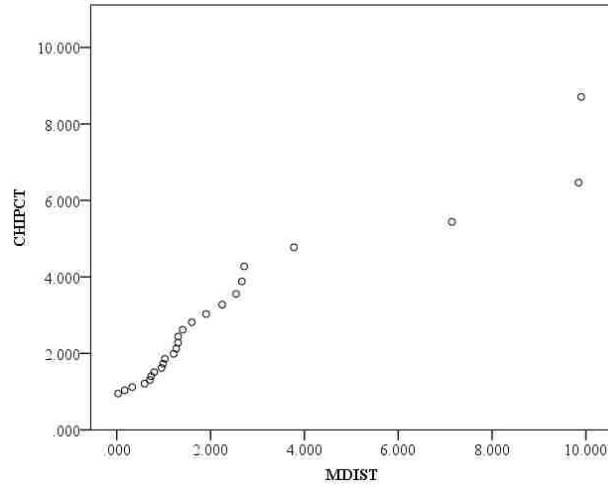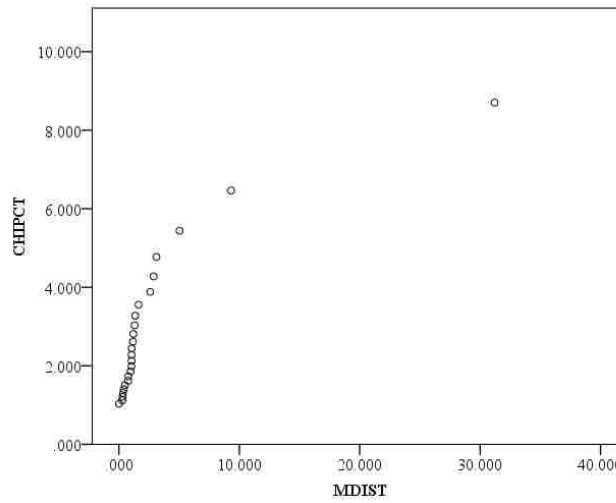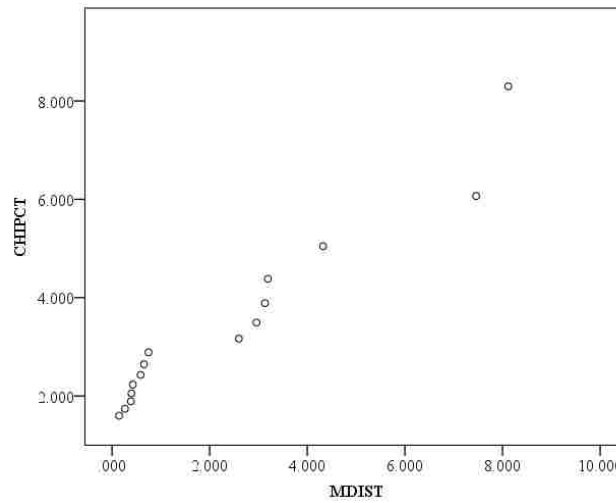
Figure 10. Graphical tests of level two multivariate normality for discipline severity models.

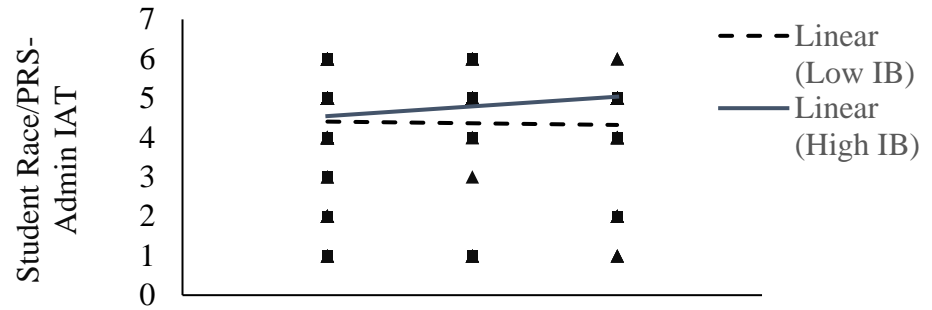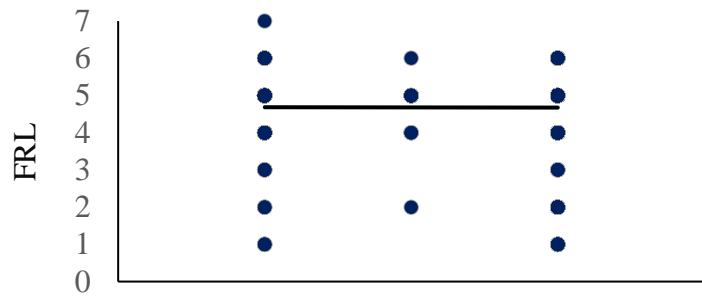*Figure 11*. Linearity assumption testing for non-significant effects in subjective discipline severity model.
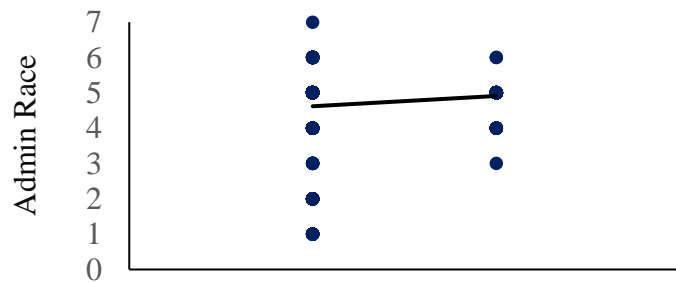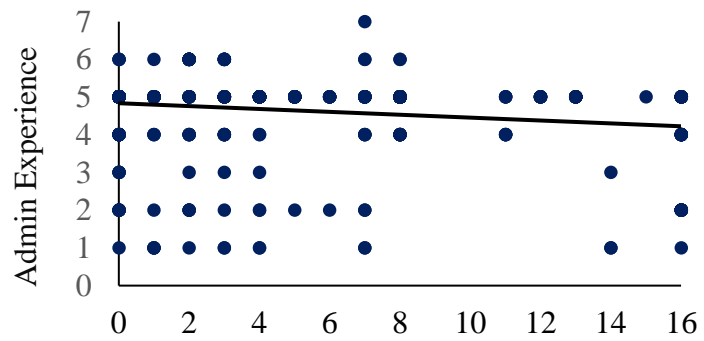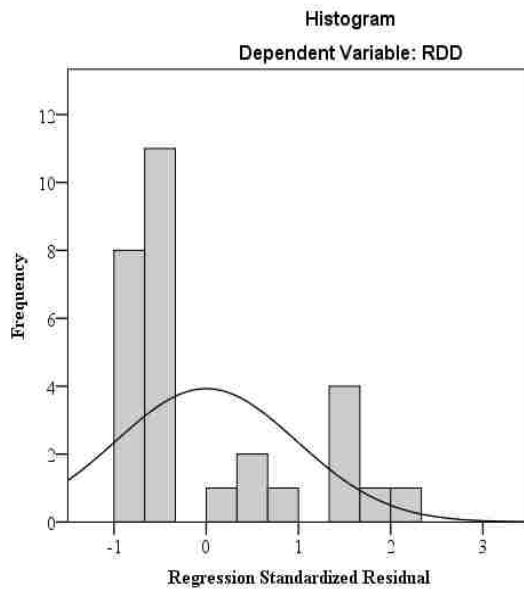
Figure 12. Linearity assumption testing for non-significant effects in objective discipline severity model
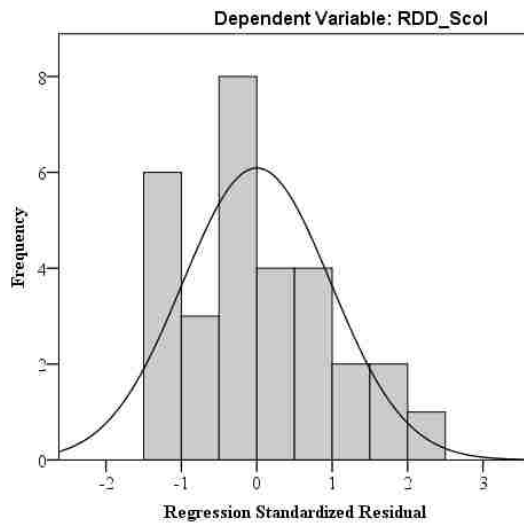
Histogram   **Normal P-P Plot**

All Discipline

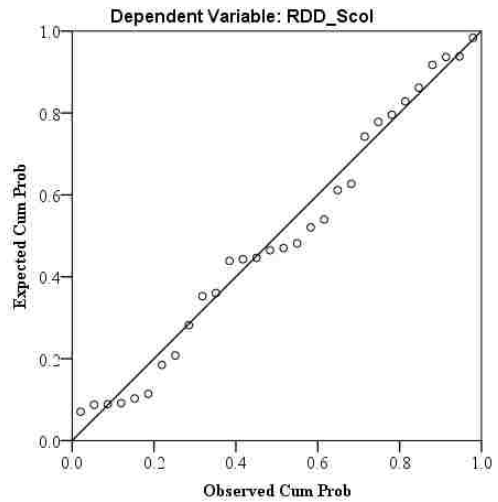**Subjective Discipline**

Figure 13. Demonstrated violations of residual normality for both discipline types when tested as

RDD in regression models.

146

REFERENCES

Alexander, C., & Arday, J. *Aiming higher: Race, inequality and diversity in the academy.*
London, UK: Runnymede.

American Bar Association, & National Bar Association. (2001). *Justice by gender: A lack of appropriate prevention, diversion, and treatment alternatives for girls in the justice system.* Retrieved from http://www.americanbar.org/content/dam/aba/publishing/criminal _justice_section_newsletter/crimjust_juvjus_justicebygenderweb.authcheckdam.pdf

Anderson, K. P., & Ritter, G. W. (2015). *Discipline disproportionalities in schools: The relationship between student characteristics and school disciplinary outcomes.* Fayetteville, AR: The University of Arkansas, Department of Education Reform.

Anyon, et al. (2014). The persistent effect of race and the promise of alternatives to suspension in school discipline outcomes. *Children and Youth Services Review, 44*, 379-386.

Balfanz, R., Byrnes, V., & Fox, J. (2013, January). *Sent home and put off-track: The antecedents, disproportionalities, and consequences of being suspended in the ninth grade*. Paper presented at the Closing the School Discipline Gap: Research to Practice conference, Washington, DC. Retrieved from https://civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/state-reports/sent-home-and-put-off-track-the-antecedents-disproportionalities-and-consequences-of-being-suspended-in-the-ninth-grade/balfanz-sent-home-ccrr-conf-2013.pdf

Banaji, M. (2013, Jul 19). *How to fight racial bias when it's silent and subtle* (S. Vedantam, Interviewer) [Audio File and Transcript]. Retrieved from http://kuer.org/post/how-fight-racial-bias-when-its-silent-and-subtle#stream/0

Banaji, M. R. (2001). Implicit attitudes can be measured. In H. L. Roehdinger, I. N. Nairne, & A. N. Suprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder.* (pp. 117-149). Washington, DC: American Psychology Association.

Banks, R. R., Eberhardt, J. L., & Ross, L. (2006). Discrimination and implicit bias in a racial unequal society. *California Law Review, 94*(4), 1169-1190.

Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science, 17*(1), 53-58.

Bartlett, T. (2017). Can we really measure implicit bias? Maybe not. *The Chronicle Review. 41*(1). Retrieved from http://www.chronicle.com/article/Can-We-Really-Measure-Implicit/238807

Beachum, F. D., & McCray, C. R. (2011). *Cultural collision and collusion: Reflections on hip-hop culture, values, and schools.* New York, NY: Peter Lang.

Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review, 6,* 242–261.

Blair, I. V., Judd, C. M., & Chapleau, C. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science, 15*(9), 674-679.

Blair, I. V., Steiner, J. F., Fairclough, D. L., Hanratty, R., Price, D. W., Hirsch, H. K., ... Havrenek, E. P. (2013). Clinicians' implicit ethnic/racial bias and perceptions of care among black and Latino patients. *Annals of Family Medicine, 11*(1), 43-52.

Blake, J. J., Smith, D. M., Marchbanks III, M. P., Seibert, A. L., Wood, S. M., & Kim, E. S. (2016). Does student–teacher racial/ethnic match impact Black students' discipline risk? a test of the cultural synchrony hypothesis. In R. J. Skiba, K. Mediratta, & M. K. Rausch (Eds.), *Inequality in School Discipline* (pp. 79-98). New York, NY: Palgrave Macmillan US.

Blanton, H., & Jaccard, J. (2006). Arbitrary metrics in psychology. *American Psychologist, 61*(1), 27-41.

Blanton, H., & Jaccard, J. (2015). Not so fast: Ten challenges to importing implicit attitude measures to media psychology. *Media Psychology, 18*(3), 338-369.

Blanton, H., Jaccard, J., & Burrows, C. (2015). Implications of the Implicit Associations Test D-transformation for psychological assessment. *Assessment, 22*(4), 429-440.

Bosson, J. K., Swann Jr, W. B., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited?. *Journal of Personality and Social Psychology, 79*(4), 631.

Bradshaw, C. P., Mitchell, M. M., O'Brennan, L. M., & Leaf, P. J. (2010). Multilevel exploration of factors contributing to the overrepresentation of Black students in office disciplinary referrals. *Journal of Educational Psychology, 102*, 508–520.

Capers, Q., Clinchot, D., McDougle, L., & Greenwald, A. G. (2017). Implicit bias in medical school admissions. *Academic Medicine, 92*(3), 365-369.

Carmichael, D., Whitten, G., & Voloudakis, M. (2005). *Study of minority over-representation in the Texas juvenile justice system.* Retrieved from http://dmcfinalreport.tamu.edu/ DMRFinalReport.pdf

Carter, P., Skiba, R., Arredondo, M., & Pollock, M. (2014). You can't fix what you don't' look at: Acknowledging race in addressing racial discipline disparities. *Discipline Disparities: A Research-to-Practice Collaborative.* Retrieved from http://youthjusticenc.org/download/education-justice/disparities/Acknowledging-Race_121514-2.pdf

Cate, I. M. P., Krolak-Schwerdt, S., & Glock, S. (2015). Accuracy of teachers' tracking

decisions: Short- and long-term effects of accountability. *European Journal of*

Psychology of Education, 31(2), 225-243.

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), 509-526.

Civil Rights Data Collection (CRDC). (2016*). 2013-2014 civil rights data collection: A first*

look. Retrieved from http://www2.ed.gov/about/offices/list/ocr/docs/2013-14-first-

look.pdf

Clark, P., & Zygmunt, E. (2014). A close encounter with personal bias: Pedagogical implication

for teacher education. *Journal of Negro Education, 83*(2), 147-161.

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

York, NY: Routledge.

Cohen, R. W. (2013). Reframing the problem: New institutionalism and exclusionary discipline

by school. *Journal of Educational Controversy, 7*(1), Article 6. Retrieved from

http://cedar.wwu.edu/jec/vol7/iss1/6

Correll, J., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity

to disambiguate potentially threatening individuals. *Journal of Personality and Social*

Psychology, 83(6), 1314-1329.

Costenbader, V., & Markson, S. (1998). School suspension: A study with secondary school

students. *Journal of School Psychology, 36*(1), 59-82.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures:

Consistency, stability, and convergent validity. *Psychological Science, 12*(2), 163-170.

Dasgupta, N., & Asgari, S. (2004) Seeing is believing: Exposure to countersterotypic women

    leaders and its effect on the malleability of automatic gender stereotyping. *Journal of*

    Experimental Social Psychology, 40(5), 642-658.

Dasgupta, N., McGhee, D. E., Greenwald, A. E., & Banaji, M. R. (2000). Automatic preference

    for White Americans: Eliminating the familiarity explanation. *Journal of Experimental*

    Social Psychology, 36, 316-328.

Devine, P. D., Forscher, P. F., Austin, A. J., & Cox, W. T. (2012). Long-term reduction in

    implicit race bias: A prejudice habit breaking. *Journal of Experimental Social*

    Psychology, 48(6), 1267-1278.

Dovidio, J. F., Kawakami, K., & Beach, K. R. (2003). Implicit and explicit attitudes:

    Examination of the relationships between measures of intergroup bias. In R. Brown & S.

    L. Gaertner (Eds.), *Blackwell handbook of social psychology: Intergroup processes*. (pp.

    175-197). Malden, MA: Blackwell Publishers.

Fabelo, T., Thompson, M. D., Plotkin, M., Carmichael, D., Marchbanks, M. P., & Booth, E. A.

    (2011). *Breaking school 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. Retrieved from

    http://www.edweek.org/media/breakingschoolsrules-37story.pdf

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*

    Power 3.1: Tests for correlation and regression analyses. *Behavior Research*

    Methods, *41*(4), 1149-1160.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic

    activation as an unobtrusive measure of racial attitudes. A bona fide pipeline*?. Journal of*

    Personality and Social Psychology, 69, 1013-1027.

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

Fenning, P., Wilczynski, J., & Parraga, M. (2000). A comparative analysis of existing secondary school discipline policies: Implication for improving practice and school safety. In K. McClafferty, C. A. Torres, & T. R. Mitchell (Eds.), *challenges of urban education: sociological perspectives for the next century* (pp. 175-194). Albany, NY: State University of New York Press.

Field, A. (2013). *Discovering statistics using IBM SPSS Statistics* (4th ed.). Thousand Oaks, CA: Sage.

Finn, J. D., & Servoss, T. J. (2015). Security measures and discipline in America high schools. In Losen, D.J., (Ed.). *Closing the school discipline gap: Equitable remedies for excessive exclusion* (pp. 44-58). New York: Teachers College Press.

FitzGerald, C., & Hurst, S. (2017). Implicit bias in healthcare professionals: A systematic review. *BMC Medical Ethics, 18:19,* 1-18.

Flanagain, W. C. (2007). *The negative aspects of in and out of school suspension and alternatives that promote academic achievement.* [doctoral dissertation] Marygrove College. Retrieved from http://www.eric.ed.gov/PDFS/ED499538.pdf

Forsyth, C. J., Biggar, R. W., Forsyth, Y. A., & Howat, H. (2014). The punishment gap: Racial/ethnic comparisons in school infractions by objective and subjective definitions. *Deviant Behavior, 36*(4), 276-287.

Gershenson, S., Holt, S. B., & Papageorge, N. (2016). Who believes in me? The effect of student-teacher demographic match on teacher expectations. *Economics of Education Review, 52,* 209-224.

Gilliam, W. S., Maupin, A. N., Reyes, C. R., Accavitti, M., & Shic, F. (2016). Do early

educators' implicit biases regarding sex and race relate to behavior expectations and

recommendations of preschool expulsions and suspensions? *Yale University Child Study*

Center. Retrieved from

http://ziglercenter.yale.edu/publications/Preschool%20Implicit%20Bias%20Policy%20Br

ief_final_9_26_276766_5379.pdf

Girvin, E. J., Gion, C., McIntosh, K., & Smolkowski, K. (2016). The relative contribution of

subjective office referrals to racial disproportionality in school discipline. *School*

Psychology Quarterly Retrieved from

https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID2796646_code1149318.pdf?abstracti

d=2796646&mirid=1

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., Kneer, J., & Kovacs, C. (2013). Preservice teachers' implicit attitudes toward students

with and without immigrations background: A piolet study. *Studies in Educational*

Evaluation, 39(4), 204-210.

Gonzalez, T. (2015). Socializing schools: Addressing racial disparities in discipline through

restorative justice. In Losen, D.J., (Ed.). *Closing the school discipline gap: Equitable*

remedies for excessive exclusion (pp. 151-165). New York NY,: Teachers College Press.

Greenwald, A. G., & Krieger, L. H. (2006). Implicit bias: Scientific foundations. *California Law*

Review, *94*(4), 945-967.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in
implicit cognition: The implicit associations test. *Journal of Personality and Social
Psychology, 74*(6), 1464-1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit
association test: I. An improved scoring algorithm. *Journal of Personality and Social
Psychology, 85*(2), 197-216.

Greenwald, A. G., Oakes, M. A., & Hoffan, H. G. (2003). Targets of discrimination: Effects of
race on responses to weapon holders. *Journal of Experimental Psychology, 39*, 399-405.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and
using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of
Personality and Social Psychology, 97*(1), 17.

Greenwald, A., Banaji, M. & Nosek, B. (2015). Statistically small effects of the Implicit
Associations Test can have large societally effects. *Journal of Personality and Social
Psychology, 108*(4), 533-561.

Greenwald, A., Nosek, B., & Sriram, N. (2006). Consequential validity of the implicit
association test: Comment on Blanton and Jaccard (2006), *American Psychologist, 61*(1),
56-61.

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.

Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A
meta-analysis. *Public Opinion Quarterly, 72*(2), 167-189.

Hatchett, S., & Schuman, H. (1976). White respondents and race of interviewer effects. *Public
Opinion Quarterly, 39*(4), 523-528.

Heilman, M. E., Wallen, A. S., Fuchs, D., & Tamkins, M. M. (2004). Penalties for success: Reactions to women who succeed at male gender-typed tasks. *Journal of Applied Psychology, 89*(3), 416-427.

Henry, P. J. (2010). Modern racism. In J. Levine & M. Hogg (Eds.), *The encyclopedia of group processes and intergroup relations* (pp. 575-577). Thousand Oaks, CA: Sage Publications.

Henry, P. J., & Sears, D.O. (2002). The symbolic racism 2000 scale. *Political Psychology, 23,* 253-283.

Jolls, C., & Sunstein, C. R. (2006). The law of implicit bias. *California Law Review, 94*(4), 969-996.

Jost, J. T., Rudman, L. A., Blair, I. V., Carney, D. R., Dasgupta, N., Glaser, J., & Hardin, C. D. (2009). The existence of implicit bias is beyond reasonable doubt: A refutation of ideological and methodological objections. *Research in Organizational Behavior, 29,* 39-69.

Kahn, K. B., & Davies, P. G. (2011). Differentially dangerous? Phenotypic racial stereotypicality increases implicit bias among ingroup and outgroup members. *Group Processes and Intergroup Relations, 14*(4), 569-580.

Kahn, K. B., Goff, P. A., & Glaser, J. (2016). Research and training to mitigate the effects of implicit stereotypes and masculinity threat on authority figures' interactions with adolescents and non-Whites. In R. J. Skiba, K. Mediratta, & M. K. Rausch (Eds.), *Inequality in school discipline: Research and practice to reduce disparities* (pp. 189-205). New York, NY: Palgrave MacMillan.

Kelly, S. (2010). A crisis in authority in predominantly Black schools? *Teachers College Record, 112*(5), 1247–1274.

Kennedy, B. L., Murphy, A. S., & Jordan, A. (2017). Title I middle school administrators'

    beliefs and choices about using corporal punishment and exclusionary

    discipline. *American Journal of Education*, *123*(2), 243-280.

Kreft, I. (1996). *Are multilevel techniques necessary? An overview, including simulation studies.*

    Los Angeles, CA: California State University.

Kreft, I., & De Leeuw, J. (1998). *Introducing multilevel modeling*. London, UK: Sage

    Publications.

Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. E. L. Joy-Gaba, J. A., … Nosek, B. A.

    (2014). Reducing implicit racial preferences: I. A comparative investigation of 17

    interventions. *Journal of Experimental Psychology: General, 143*(4), 1765‑1785.

Lee, C. (2013). Making race salient: Trayvon Martin and implicit bias in not yet a post-racial

    society. *North Carolina Law Review, 91,* 101-157.

Levinson, J. D. (2007). Forgotten racial equality: Implicit bias, decision making, and

    misremembering. *Duke Law Journal, 57*(2), 345-424.

Lopez, G. (2017, Mar 7). *For years, this popular test measured anyone's racial bias. But it might

    not work after all.* Retrieved from: http://www.vox.com/identities/2017/3/7/14637626/

    implicit-association-test-racism

Losen, D. J., & Skiba, R. J. (2010). *Suspended education: Urban middle schools in crisis.*

    Southern Poverty Law Center. Retrieved from

    https://www.splcenter.org/sites/default/files/d6_legacy_files/downloads/publication/Susp

    ended_Education.pdf

Losen, D. J., Hodson, C. L., Keith, I. I., Michael, A., Morrison, K., & Belway, S. (2015). *Are we

    closing the school discipline gap? K-12 racial disparities in school discipline.* The Center

    for Civil Rights Remedies. Retrieved from

https://civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-

remedies/school-to-prison-folder/federal-reports/are-we-closing-the-school-discipline-

gap/AreWeClosingTheSchoolDisciplineGap_FINAL221.pdf

Losen, D., & Haynes, L. S. E. (2016). Eliminating excessive and disparate school discipline: A

review of research and policy reform. In R. J. Skiba, K. Mediratta, & M. K. Rausch

(Eds.) *Inequality in school discipline* (pp. 245-262). New York, NY: Palgrave Macmillan.

Losen, D., Sun, W., & Keith, M. A. (2017). Suspended education in Massachusetts: Using days

of lost instructions due to suspension to evaluate our schools. *The Center for Civil Rights

Remedies.* Retrieved from

https://www.civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-

remedies/school-to-prison-folder/summary-reports/suspended-education-in-

massachusetts-using-days-of-lost-instruction-due-to-suspension-to-evaluate-our-

schools/suspended-education-web-corrected-032317.pdf

Lueke, A., & Gibson, B. (2015). Mindfulness meditation reduces implicit age and race bias: The

role of reduced automaticity of responding. *Social Psychological and Personality

Science, 6*(3), 284‑291.

Maas, C. J., & Hox, J. J. (2002). Robustness of multilevel parameter estimates against small

sample sizes. *Unpublished Paper, Utrecht University*. Retrieved from

https://www.researchgate.net/profile/Joop_Hox/publication/228576801_Robustness_of_

multilevel_parameter_estimates_against_small_sample_sizes/links/0046352975308ef5cf

000000.pdf

Maddox, K. B. (2004). Perspective on racial phenotypicality bias. *Personality and Social

Psychology Review, 8*(4), 383-401.

Maddox, K. B., & Gray, S., A. (2002). Cognitive representation of Black Americans: Re-exploring the role of skin tone. *Personality and Social Psychology Bulletin, 28*(2), 250-259.

Mallett, C. A. (2016). The school-to-prison pipeline: A critical review of the punitive paradigm shift. *Child and Adolescent Social Work Journal*, *33*(1), 15-24.

Mann, T. C., & Ferguson, M. J. (2015). Can we undo our first impressions? The role of reinterpretation in reversing implicit evaluations. *Journal of Personality and Social Psychology, 108*(6), 823‑849.

Marchbanks, M. P., Blake, J. J., Booth, E. A., Carmichael, D., Seibert, A., & Fabelo, T. (2015) The economic effects of exclusionary discipline on grade retention and high school dropout. In Losen, D. J. (Ed.), *Closing the school discipline gap: Equitable remedies for excessive exclusion* (pp. 59-74). New York, NY: Teachers College Press.

Marchbanks, M. P., Blake, J. J., Smith, D., Seibert, A. L., & Carmichael, D. (2014). More than a drop in the bucket: The social and economic costs of dropouts and grade retentions associated with exclusionary discipline. *Journal of Applied Research on Children, 5*(2), 1-34.

McConahay, J. B. (1986). Modern racism, ambivalence, and the modern racism scale. In J. F. Dovidio, & S. L. Gaertner (Eds.), *Prejudice, discriminations, and racism* (pp. 91-125). San Diego, CA: Academic Press.

McFadden, A. C., Marsh II, 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.

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), 1-21.

Milkman, K. L., Akinola, M., & Chugh, D. (2015). What happens before? A field experiment exploring how pay and representation differentially shape bias on the pathway into organizations. *Journal of Applied Psychology, 100*(6), 1678-1712.

Molenberghs, P. (2013). The neuroscience of in-group bias. *Neuroscience & Biobehavioral Reviews*, *37*(8), 1530-1536.

Monroe, C. R. (2005). Understanding the discipline gap through a cultural lens: Implications for the education of African American students. *Intercultural Education, 16*(4), 317-330.

Morris, E. W., & Perry, B. L. (2016). The punishment gap: School suspension and racial disparities in achievement. *Social Problems, 63*(1), 68-86.

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

Muthen, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological methodology, 25*, 267-316.

Nance, J. P. (2016). Over-disciplining students, racial bias, and the school-to-prison-pipeline. *University of Florida Law Scholarship Repository.* Retrieved from http://scholarship.law.ufl.edu/cgi/viewcontent.cgi?article=1766&context=facultypub

Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied linear statistical models* (Vol. 4). Chicago, IL: Irwin.

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

Nishioka, V. (with Shigeoka, S., & Lolich, E.). (2017). *School discipline data indicators: A guide for districts and schools* (REL 2017–240). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and

Regional Assistance, Regional Educational Laboratory Northwest. Retrieved from http://ies.ed.gov/ncee/edlabs.

Noltemeyer, A. L., & Mcloughlin, C. S. (2010a). Changes in exclusionary discipline rates and disciplinary disproportionality over time. *International Journal of Special Education, 25*(1), 59-70.

Noltemeyer, A., & Mcloughlin, C. S. (2010b). Patterns of exclusionary discipline by school typology, ethnicity, and their interaction. *Perspectives on Urban Education*, 27-40.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The implicit association test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes* (pp. 265–292). New York: Psychology Press.

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., … & Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology, 18,* 36 – 88.

Obama, B. (2015). *Eulogy for the Honorable Reverend Clementa Pinckney* (speech transcript). Retrieved from https://www.whitehouse.gov/the-press-office/2015/06/26/remarks-president-eulogy-honorable-reverend-clementa-pinckney

Office of Civil Rights. (2011-12). *District and school report.* Available at http://ocrdata.ed.gov/DistrictSchoolSearch

Office of Civil Rights. (2014). *Civil rights data collection: Data snapshot (school discipline).* Washington, DC: US Department of Education Office for Civil Rights. Retrieved from http://www2.ed.gov/about/offices/list/ocr/docs/crdc-discipline-snapshot.pdf

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

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2015). Using the IAT to

      predict ethnic and racial discrimination: Small effect sizes of unknown societal

      significance. *Journal of Personality and Social Psychology, 108*(4), 562-571.

Perry, B. L., & Morris, E. W. (2014). Suspending progress: Collateral consequences of

      exclusionary punishment in public schools. *American Sociological Review,* 1-21.

Project Implicit. (2016, February 21). *Project implicit education: FAQ.* Retrieved from Project

      Implicit Web site: https://implicit.harvard.edu/implicit/faqs.html#faq10

Pudelski, S. (2014). *School discipline in the eyes of school superintendents.* AASA and The

      Children's Defense Fund. Retrieved from

      http://www.aasa.org/uploadedFiles/Childrens_Programs/_files/SurveyReport-School-

      Discipline-in-the-Eyes-of-Superintendents.pdf

Rafaelle 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.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data

      analysis methods* (2nd edition). Thousand Oaks, CA: Sage.

Raudenbush, S.W., Bryk, A.S, & Congdon, R. (2013). HLM 7.01 for Windows [Computer

      software]. Skokie, IL: Scientific Software International, Inc.

Rausch, M. K., & Skiba, R. (2004*). Disproportionality in school discipline among minority

      students in Indiana: Description and analysis.* Bloomington, IN: Center for Evaluation

      and Education Policy.

Rausch, M. K., & Skiba, R. J. (2005). *The academic cost of discipline: The relationship between

      suspension/expulsion and school achievement.* Annual meeting of the American

      Educational Research Association, (pp. 1-43). Montreal, Canada.

Reise, S. P., & Duan, N. (2003). *Multilevel modeling: Methodological advances, issues, and applications.* New York, NY: Routldge.

Relethford, J. H. (2009). Race and global patterns of phenotypic variation. *American Journal of Physical Anthropology, 139*(1), 16-22.

Reschly, D. J. (1997). *Disproportionate minority representation in general and special education: Patterns, issues, and alternatives.* Des Moines, IA: Mountain Plains Regional Resource Center, Drake University.

Ronquillo, J., Denson, T. F., Lickel, B., Lu, Z., Nandy, A., & Maddox, K. (2007). The effects of skin tone on race-related amygdala activity: An fMRI investigation. *Social Cognitive and Affective Neuroscience, 2*(1), 39-44.

Rudman, L. A., & Ashmore, R. D. (2007). Discrimination and the implicit associations test. *Group Processes and Intergroup Relations, 10*, 359-372.

Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues, 57*, 743-762.

Rumberger, R. W., & Losen, D. (2017). The hidden cost of California's harsh school discipline and the localized economic benefits from suspending fewer high school students. *The Center for Civil Rights Remedies.* Retrieved from https://www.civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-remedies/school-to-prison-folder/summary-reports/the-hidden-cost-of-californias-harsh-discipline/CostofSuspensionReportFinal-corrected-030917.pdf

Rumberger, R. W., & Losen, D. J. (2016). *The high cost of harsh discipline and its disparate impact*. The Center for Civil Rights Remedies. Retrieved from https://www.civilrightsproject.ucla.edu/resources/projects/center-for-civil-rights-

remedies/school-to-prison-folder/federal-reports/the-high-cost-of-harsh-discipline-and-its-disparate-impact/UCLA_HighCost_6-2_948.pdf

Safe Supportive Learning (2016). *Pennsylvania compilation of school discipline laws and regulations.* Retrieved from https://safesupportivelearning.ed.gov/sites/default/files/discipline-compendium/Pennsylvania%20School%20Discipline%20Laws%20and%20Regulations.pdf

Safer, D. J. (1986). Nonpromotion correlates and outcomes at different grade levels. *Journal of Learning Disabilities, 19*(8), 500-503.

Safer, D. J., Heaton, R. C., & Parker, F. C. (1981). A behavioral program for disruptive junior high school students: Results and follow-up. *Journal of Abnormal Child Psychology, 9*(4), 483-494.

Schuman, H., Steeh, C., Bobo, L., Krysan, M. (1997). *Racial attitudes in America: Trends and interpretations.* Cambridge, MA: Harvard University Press.

Schwartz, R. H., & Wirtz, P. W. (1990). Potential substance abuse detection among adolescent patients. *Clinical Pediatrics, 29*(1), 38-43.

Scott, M., & Barrett, S. B. (2004). Using staff and student time engaged in disciplinary procedures to evaluate the impact of schoolwide PBS. *Journal of Positive Behavior Interventions, 6*(1), 21-27.

Shollenberger, T.L. (2015). Racial disparities in school suspension and subsequent outcomes: Evidence from the national longitudinal survey of youth. In Losen, D.J. (Ed.), *Closing the school discipline gap: Equitable remedies for excessive exclusion* (pp. 31-43). New York, NY: Teachers College Press.

Simson, D. (2014). Exclusion, punishment, racism, and our schools: A critical race theory

perspective on school discipline. *UCLA Law Review, 62*(3), 506-565.

Singer, J. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and

individual growth models. *Journal of Education and Behavioral Statistics, 24*(4), 323-

355.

Skiba, R. J., & Edl, H. M. (2004). The disciplinary practices survey: How do Indiana's principals

feel about discipline?. *Children Left Behind Policy Briefs. Supplementary Analysis*.

Skiba, R. J., & Rausch, M. K. (2004). The relationship between achievement, discipline, and

race: An analysis of factors predicting ISTEP scores. *Children Left Behind Policy Briefs*,

1-5.

Skiba, R. J., & Rausch, M. K. (2006). Zero tolerance, suspension, and expulsion: Questions of

equity and effectiveness. In C. M. Evertson, & C. M. Weinstein (Eds.), *Handbook of

classroom management: Research, practice, and contemporary issues* (pp. 1063-1089).

Mahwah, NJ: Lawrence Erlbaum Associates.

Skiba, R. J., & Williams, N. T. (2014). Are Black kids worse? Myths and facts about racial

differences in behavior. *The Equity Project at Indiana University.* Retrieved from

http://www.indiana.edu/~atlantic/wp-content/uploads/2014/03/African-American-

Differential-Behavior_031214.pdf

Skiba, R. J., Chung, C. G., Trachok, M., Baker, T. L., Sheya, A., & Hughes, R. L. (2014).

Parsing disciplinary disproportionality: Contribution of infraction, student, and school

characteristics to out-of-school suspension and expulsion. *American Educational

Research Journal, 51*(4), 640-670.

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

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

Skiba, R. J., Shure, L., & Williams, N. (2012). Racial and ethnic disproportionality in suspension and expulsion. In A. L. Noltemeyer & C. S. Mcloughlin (Eds.), *Disproportionality in educational and special education* (pp. 89-118). Springfield, IL: Charles C. Thomas Publisher.

Skiba, R. J., Simmons, A., Staudinger, L., Rausch, M., Dow, G., & Feggins, R. (2003). *Consistent removal: Contributions of school discipline.* Paper presented at School-to-Prison-Pipeline Conference: Harvard Civil Rights Project.

Skiba, R., J., Trachok, M., Chung, C. G., Baker, T., Sheya, A., & Hughes, R. (2013, January). *Where should we intervene? Contributions of behavior, student, and school characteristics to suspension and expulsion*. Paper presented at the Closing the School Discipline Gap: Research to Practice conference, Washington, DC.

Skiba, R.J., Mediratta, K., & Rausch, M. K. (2016). *Inequality in school discipline*. New York, NY: Palgrave Macmillan US.

Skrla, L., Scheurich, J. J., Garcia, J., & Nolly, G. (2004). Equity audits: A practical leadership tool for developing equitable and excellent schools. *Educational Administration Quarterly, 40*(1), 133-161.

Smolkowski, K., Girvan, E. J., McIntosh, K., Nese, R. & Horner, R. H. (2017). Vulnerable decision points for disproportionate office discipline referrals: Comparisons of discipline

for African American and White elementary school students. *Behavioral Disorders, 41*(4), 178-195.

Snijders, T. A. B., & Bosker, R. J. (1993). Standard errors and sample sizes for two-level research. *Journal of Educational Statistics, 18*(3), 237-259.

Snijders, T. A. B., & Bosker, R. J. (1994). Modeled variance in two-level models. *Sociological Methods Research, 22*, 342-363.

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London, UK: Sage Publishers.

Solórzano, D. G., & Yosso, T. J. (2002). Critical race methodology: Counter-storytelling as an analytical framework for educational research. *Qualitative Inquiry, 8*(1), 23-44.

Staats, C. (2015). Understanding implicit bias: What educators should know. *American Educator, 2015*(Winter). Retrieved from http://www.aft.org/ae/winter2015-2016/staats

Staats, C., Capatosto, K., Wright, R. A., & Jackson, V. W. (2016). *2016 State of the science: Implicit bias review*. Columbus, OH: Kirwan Institute for the Study of Race and Ethnicity. Retrieved from http://kirwaninstitute.osu.edu/wp-content/uploads/2016/07/implicit-bias-2016.pdf

Stanley, D. A., Sokol-Hessner, P., Banaji, M. R., & Phelps, E. A. (2011). Implicit race attitude predicts trustworthiness judgements and economic trust decisions. *Proceedings of the National Academy of Sciences, 108*(19), 7710-7715.

Steinberg, M. P., Allensworth, E., & Johnson, D. W. (2015). What conditions support safety in urban schools? The influence of school organizational practices on student and teacher reports of safety in Chicago. In D. J. Losen (Ed.), *Closing the school discipline gap: Equitable remedies for excessive exclusion* (pp. 118-131). New York, NY: Teachers College Press.

Stell, A. J., & Farsides, T. (2015). Brief loving-kindness meditation reduces racial bias, mediated by positive other-regarding emotions. *Motivation and Emotion*, Retrieved from https://link.springer.com/article/10.1007/s11031-015-9514-x

Stoet, G. (2010). PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods, 42*(4), 1096-1104.

Stoet, G. (*in press*). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*.

Tetlock, P. E., & Mitchell, G. (2009). Implicit bias and accountability systems: what must organizations do to prevent discrimination? *Research in Organizational Behavior, 29*, 3-38.

Texas Appleseed. (2007). *Texas' school-to-prison-pipeline - dropout to incarceration: The impact of school discipline and zero tolerance.* Austin, TX: Texas Appleseed.

The Civil Rights/Advancement Project. (2000). O*pportunities suspended: The devastating consequences of zero tolerance and school discipline*. Boston, MA: Harvard University.

The Leadership Conference Education Fund. (2016). *Every student succeeds action: Guide for advocates.* Washington, DC. Retrieved from http://civilrightsdocs.info/pdf/education/ESSA/ESSA-Guide.pdf

Tobin, T., Sugai, G., & Colvin, G. (1996). Patterns in middle school discipline records. *Journal of Emotional and Behavioral Disorders, 4*(2), 82-94.

Townsend, B. L. (2000). The disproportionate discipline of African American learners: Reducing school suspensions and expulsions. *Exceptional Children, 66*(2), 381-391.

US General Accounting Office. (1990). *Death penalty sentencing: Research indicates pattern of racial disparities.* Report to the Senate and House Committees on the Judiciary, Washington, DC. Retrieved from http://www.gao.gov/assets/220/212180.pdf

van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M., & Holland, R. W. (2010). The implicit prejudiced attitude of teachers: Relations to teacher expectations and the ethnic achievement gap. *American Educational Research Journal, 47,* 497-527.

Vincent, C. G., Randall, C., Cartledge, G., Tobin, T. J., Swain-Bradley, J. (2011). Toward a conceptual integration of cultural responsiveness and schoolwide positive behavior support. *Journal of Positive Behavior Interventions, 13*(4), 219-229.

Wald, J. (2014). *Supplementary paper II: Can debiasing strategies help reduce racial disparities in school discipline.* Boston, MA: Charles Hamilton Houston Institute for Race & Justice (Harvard University).

Wald, J., & Losen, D. (2003). *Defining and redirecting a school-to-prison pipeline. School-to-prison pipeline research conference* (pp. 1-11). Boston, MA: The Civil Rights Project at Harvard University.

Warikoo, N., Sinclair, S., Fei, J., & Jacoby-Senghor, D. (2016). Examining racial bias in education: A new approach. *Educational Researcher, 45*(9), 508-514.

Wright, A. C. (2015). Teachers' perceptions of students' disruptive behavior: The effect of racial congruence and consequences for school suspension. [Unpublished manuscript] Santa Barbara, CA: University of California Department of Economics. Available from https://aefpweb.org/sites/default/files/webform/41/Race%20Match,%20Disruptive%20Behavior,%20and%20School%20Suspension.pdf

Wu, S.-C. (1980). *The foundations of student suspension*. Honolulu, HI: University of Hawaii.

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

Yusuf, A. R., Irvine, A., & Bell, J. (2016). Reducing racial disparities in school discipline: Structured decision-making in the classroom. In R. J. Skiba, K. Mediratta, & M. K.

Rausch (Eds.), *Inequality in School Discipline* (pp. 99-114). New York, NY: Palgrave

    Macmillan US.

Zygmunt, E., & Clark, P. (2016). *Transforming teacher preparation for social justice.* New

    York, NY: Teachers College Press.

APPENDIX A

Invitation Email

Dear [Superintendent, Principal, Vice Principal]:

My name is Gina Gullo and I am conducting my dissertation research under the supervision of Dr. Floyd Beachum in the department of Educational Leadership at Lehigh University. For this study, I am looking for schools/districts to share their discipline data including the following elements for students receiving school discipline:

- ID (no names as to coincide with FERPA regulations)
- Grade
- Free or reduced lunch eligibility
- Race/ethnicity
- behavioral infraction
- disciplinary outcome
- ID for the administrator who determined the disciplinary outcome

Furthermore, I'd like to have any administrators involved in disciplinary decisions complete a short survey including a measure of unintentional (or implicit) bias, their years of experience, and their race/ethnicity identification.

Compensation for participation will be provided in the form of a district equity audit using the provided discipline data. Administrator bias test participation is not necessary to receive this compensation. All districts and schools sharing any data will receive this anonymous analysis at the conclusion of the study. This report includes percentages and graphs to help schools know if discipline is being equitably administered in the district and if behavioral infractions are equitably occurring. An example Equity Audit report is attached.

I ask that if you choose to participate, you share your school or district data with me once in the next month and then one final time after 50% (typically 90 days) of school for that district have been completed. Additionally, all participating districts will have access to an online tool to measure implicit and explicit bias and learn about how to lessen the effects of each in educational practice after the study is completed.

Please consider participating in this study. If you decide to participate, please return the Site Permission Letter and Informed Consent via email (GLC211@Lehigh.edu), fax (610-758-3227), or mail (see address below). Feel free to call me if you have any questions or concerns.

Thank you,

Gina Gullo
Doctoral Candidate, Lehigh University
201-618-3075; GLC211@lehigh.edu

Address:
Gina Gullo
Lehigh University
Iacocca Hall A-208
111 Research Drive
Bethlehem, PA 18015

School Discipline Equity Audit

Example School District: 2016-2017

Executive Summary

> *When looking at graphs compare each bar to the black bar. The black bar represents the proportion of each group in the school/district. Bars higher or lower than the black bar indicate a discrepancy with more or less of that group experiencing the noted condition. A discrepancy is not considered inequitable until it is ten percent or greater in either direction.*

This equity audit is based on data shared with the researcher relative to discipline by the school or school district. Data revealed that overall the school/district has a _____ level of equity in relation to school discipline and a _____ level of equity in relation to school behaviors. Equity is determined by a discrepancy of less than ten percent between the representation of a group in the population (i.e. school or district) and the representation of that group on an indicator (i.e. suspensions or infractions).

- An excellent level of school discipline equity indicates students were equitably administered in-school (ISS) and out-of-school suspensions (OSS) on all factors measured (free-or-reduced lunch eligibility, gender, and race).
- A high level of equity indicates student equity on ISS and OSS on two out of three factors.
- A moderate level of equity indicates student equity on ISS and OSS on one out of three factors.
- A low level of equity indicates student equity on ISS and OSS on none of three factors.
- The same pattern exists for school behavior equity with regards to student behavior infractions on each item.

Methodology

Data received from the school/district were analyzed in Excel and graphed for ease of understanding. The collection of the data occurred on the school/district end and is not documented in this report.

172

Demographics

The current school/district has:

- ____% of students eligible for free or reduced lunch (FRL)
- ____% male students; ____% female students
- ____% White students; ____% Black/African American students; ____% Hispanic/Latino students; ____% Student of other Ethnicities
- The school/district has a total of _____ students in the population
- _____Students received an ISS during data collection
- _____Students received an OSS during data collection
- _____Students had no behavior infractions during data collection
- _____Students had only one behavior infraction during data collection
- _____Students had two or more behavior infractions during data collection

Discipline and Behavior by FRL Eligibility

- There was a discrepancy of _____% between students eligible for FRL receiving ISS and students in the population eligible for FRL with students eligible for FRL receiving (more/less) ISS than those ineligible for FRL.
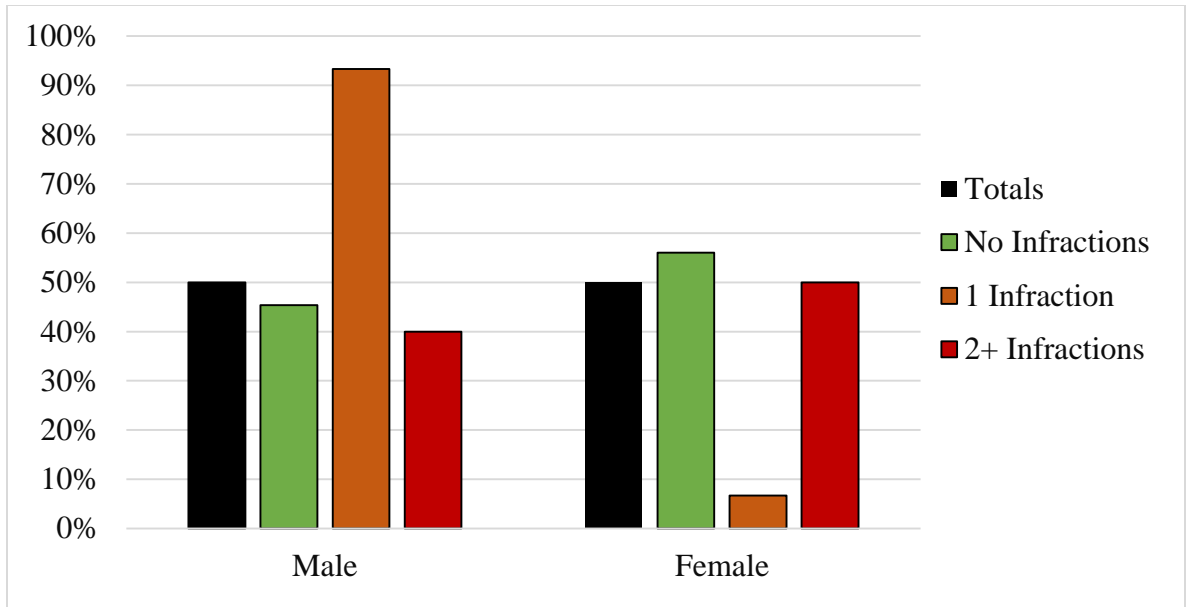- There was a discrepancy of _____% between students eligible for FRL receiving OSS and students in the population eligible for FRL with students eligible for FRL receiving (more/less) OSS than those ineligible for FRL.
- There was a discrepancy of _____% between students eligible for FRL exhibiting no behavioral incidents and students in the population eligible for FRL with students eligible for FRL exhibiting (more/less) behavioral incidents than those ineligible for FRL.
- There was a discrepancy of _____% between students eligible for FRL exhibiting only one behavioral incident and students in the population eligible for FRL with students eligible for FRL exhibiting only one behavioral incident (more/less) often than those ineligible for FRL.
- There was a discrepancy of _____% between students eligible for FRL exhibiting two or more behavioral incidents and students in the population eligible for FRL with students eligible for FRL exhibiting two or more behavioral incidents (more/less) often than those ineligible for FRL.
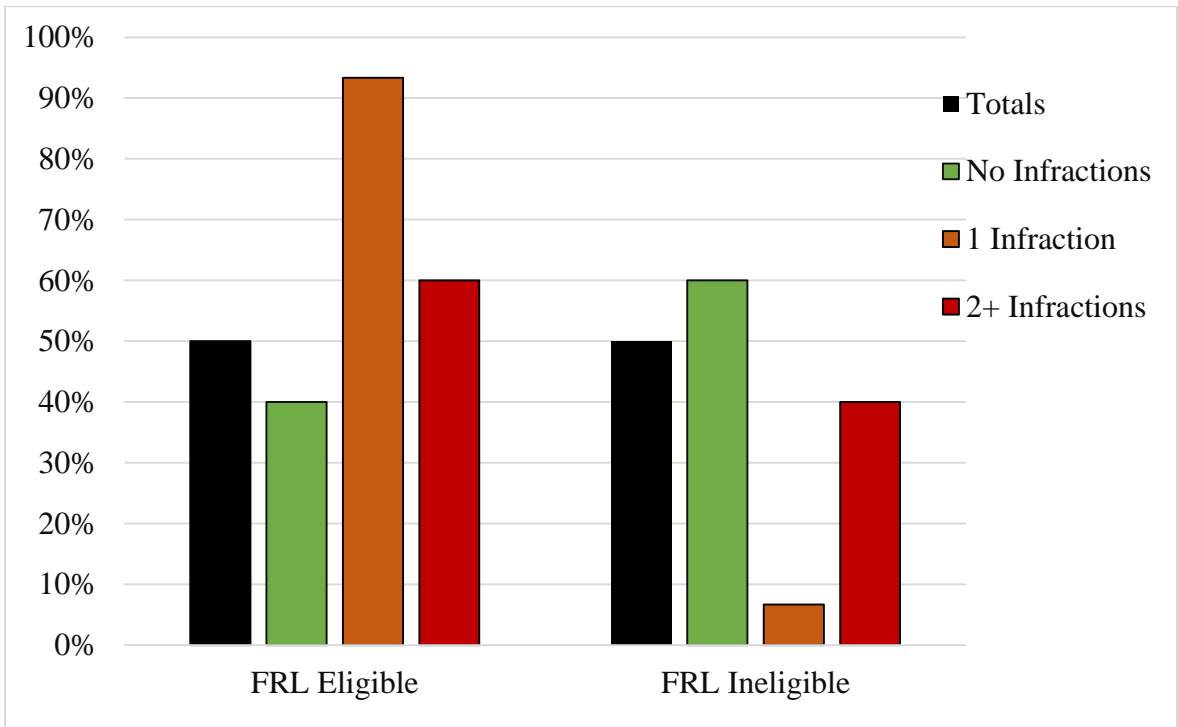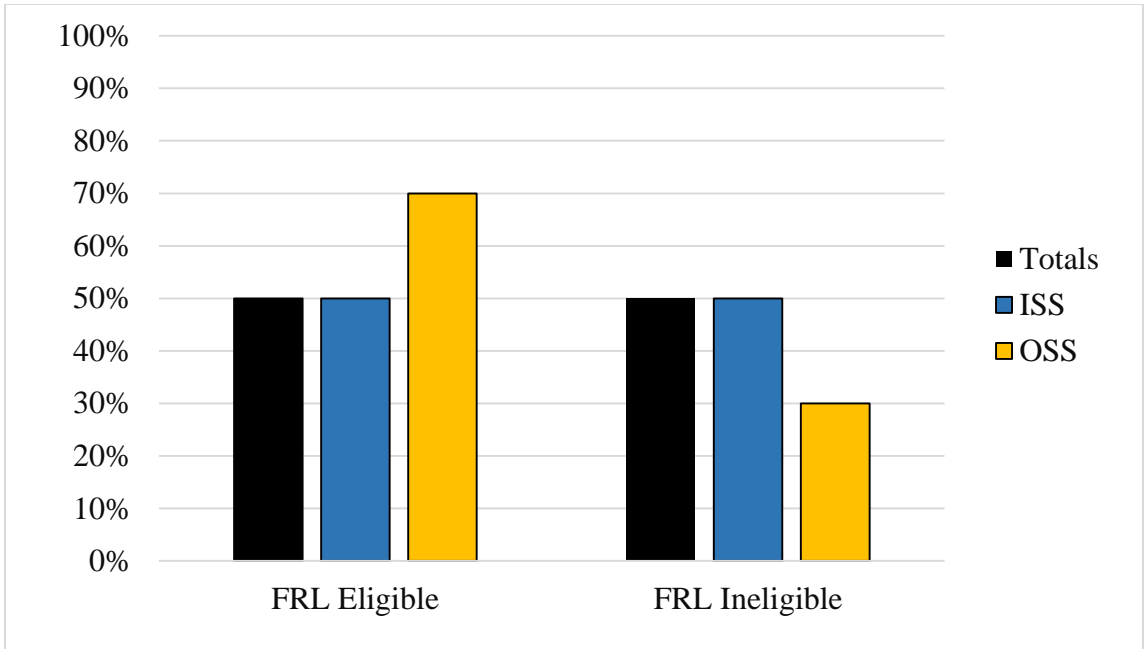
Discipline and Behavior by Gender

- There was a discrepancy of _____% between males receiving ISS and males in the population with males receiving (more/less) ISS than females.
- There was a discrepancy of _____% between males receiving OSS and males in the population with males receiving (more/less) OSS than females.
- There was a discrepancy of _____% between males exhibiting no behavioral incidents and males in the population with males exhibiting (more/less) behavioral incidents than females.
- There was a discrepancy of _____% between males exhibiting only one behavioral incident and males in the population with males exhibiting only one behavioral incident (more/less) often than females.
- There was a discrepancy of _____% between males exhibiting two or more behavioral incidents and males in the population with males exhibiting two or more behavioral incidents (more/less) often than females.
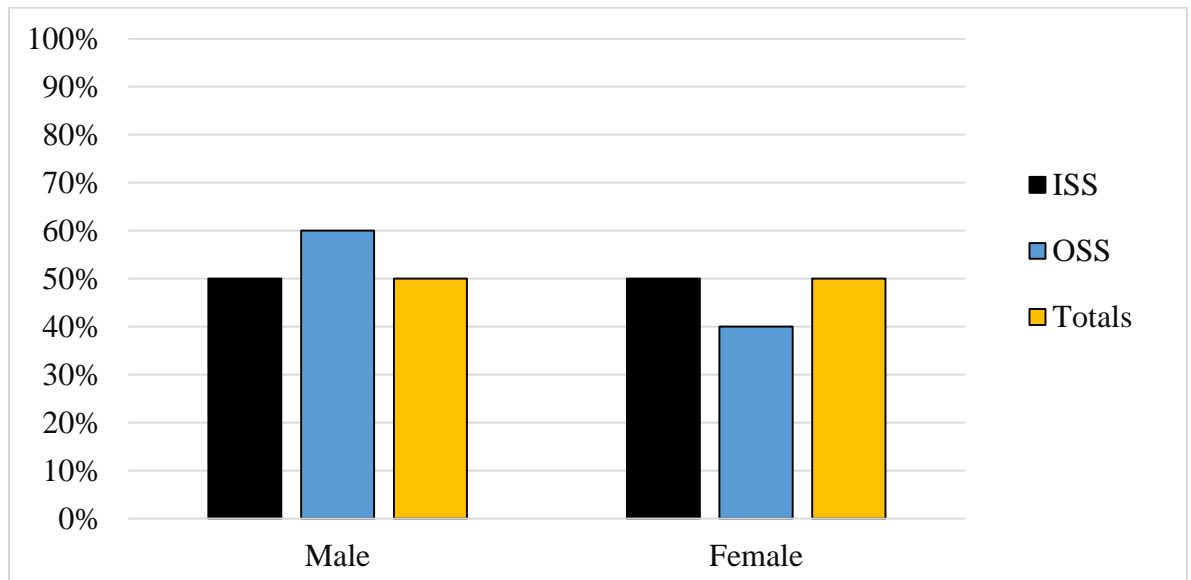
Discipline and Behavior by Race/Ethnicity

White/Caucasian

- There was a discrepancy of _____% between students identifying as White receiving ISS and White students in the population with White students receiving (more/less) ISS than students of other races/ethnicities.
- There was a discrepancy of _____% between students identifying as White receiving OSS and White students in the population with White students receiving (more/less) OSS than students of other races/ethnicities.
- There was a discrepancy of _____% between students identifying as White exhibiting no behavioral incidents and White students in the population with White students exhibiting (more/less) behavioral incidents than students of other races/ethnicities.
- There was a discrepancy of _____% between students identifying as White exhibiting only one behavioral incident and White students in the population with White students exhibiting only one behavioral incident (more/less) frequently than students of other races/ethnicities.
- There was a discrepancy of _____% between students identifying as White exhibiting two or more behavioral incidents and White students in the population with White students exhibiting two or more behavioral incidents (more/less) frequently than students of other races/ethnicities.

Black/African American

- There was a discrepancy of _____% between students identifying as Black receiving ISS and Black students in the population with Black students receiving (more/less) ISS than students of other races/ethnicities.
- There was a discrepancy of _____% between students identifying as Black receiving OSS and Black students in the population with Black students receiving (more/less) OSS than students of other races/ethnicities.
- There was a discrepancy of _____% between students identifying as Black exhibiting no behavioral incidents and Black students in the population with Black students exhibiting (more/less) behavioral incidents than students of other races/ethnicities.
- There was a discrepancy of _____% between students identifying as Black exhibiting only one behavioral incident and Black students in the population with Black students exhibiting only one behavioral incident (more/less) frequently than students of other races/ethnicities.
- There was a discrepancy of _____% between students identifying as Black exhibiting two or more behavioral incidents and Black students in the population with Black students exhibiting two or more behavioral incidents (more/less) frequently than students of other races/ethnicities.

Hispanic/Latino(a)
- There was a discrepancy of _____% between students identifying as Hispanic receiving ISS and Hispanic students in the population with Hispanic students receiving (more/less) ISS than students of other races/ethnicities.
- There was a discrepancy of _____% between students identifying as Hispanic receiving OSS and Hispanic students in the population with Hispanic students receiving (more/less) OSS than students of other races/ethnicities.
- There was a discrepancy of _____% between students identifying as Hispanic exhibiting no behavioral incidents and Hispanic students in the population with Hispanic students exhibiting (more/less) behavioral incidents than students of other races/ethnicities.
- There was a discrepancy of _____% between students identifying as Hispanic exhibiting only one behavioral incident and Hispanic students in the population with Hispanic students exhibiting only one behavioral incident (more/less) frequently than students of other races/ethnicities.
- There was a discrepancy of _____% between students identifying as Hispanic exhibiting two or more behavioral incidents and Hispanic students in the

population with Hispanic students exhibiting two or more behavioral incidents (more/less) frequently than students of other races/ethnicities.

Other Race/Ethnicities (Asian, Pacific Islander, Native American, Multiracial, No Response)

- There was a discrepancy of _____% between students identifying as Other Race/Ethnicity receiving ISS and Other Race/Ethnicity students in the population with Other Race/Ethnicity students receiving (more/less) ISS than students identifying as White, Black, or Hispanic.
- There was a discrepancy of _____% between students identifying as Other Race/Ethnicity receiving OSS and Other Race/Ethnicity students in the population with Other Race/Ethnicity students receiving (more/less) OSS than students identifying as White, Black, or Hispanic.
- There was a discrepancy of _____% between students identifying as Other Race/Ethnicity exhibiting no behavioral incidents and Other Race/Ethnicity students in the population with Other Race/Ethnicity students exhibiting (more/less) behavioral incidents than students identifying as White, Black, or Hispanic.
- There was a discrepancy of _____% between students identifying as Other Race/Ethnicity exhibiting only one behavioral incident and Other Race/Ethnicity students in the population with Other Race/Ethnicity students exhibiting only one behavioral incident (more/less) frequently than students identifying as White, Black, or Hispanic.
- There was a discrepancy of _____% between students identifying as Other Race/Ethnicity exhibiting two or more behavioral incidents and Other Race/Ethnicity students in the population with Other Race/Ethnicity students exhibiting two or more behavioral incidents (more/less) frequently than students identifying as White, Black, or Hispanic.
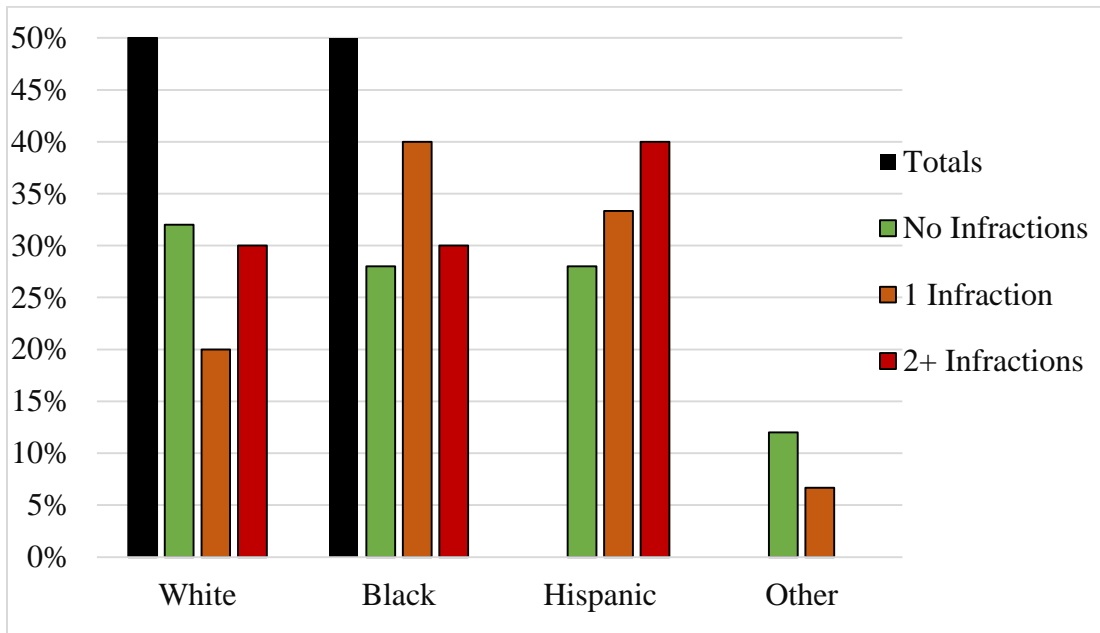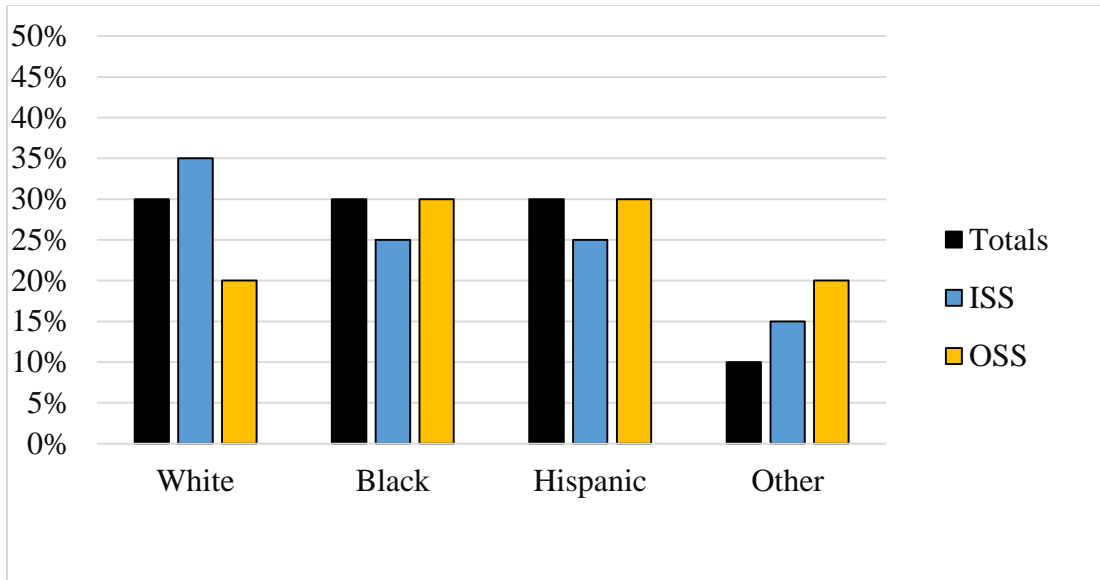
APPENDIX C

INFORMED CONSENT FORM

LEHIGH
U N I V E R S I T Y.

CONSENT FORM
Implicit Bias in School Disciplinary Decisions

You are invited to be in a research study of implicit bias as it relates to school discipline. You were selected as a possible participant because your district has a diverse student population and reported disciplinary infraction in the 2014-2015 school year. We ask that you read this form and ask any questions you may have before agreeing to be in the study.

This study is being conducted by: Gina Gullo, M.Ed., Educational Leadership (Lehigh University), under the direction of Floyd Beachum, Ed.D., Educational Leadership (Lehigh University)

Purpose of the study:
The purpose of this study is: to better understand how unintentional bias might be involved in decision making for Pennsylvania principals and vice principals when making decisions about behavior consequences for discipline infractions.

Procedures
If you agree to be in this study, we would ask you to do the following things:
Share your school discipline database with the researcher inclusive of student ID (not name), student grade; student free or reduced lunch eligibility, student race, student behavioral infraction, student disciplinary outcome, and an ID for the administrator who determined the disciplinary outcome. The administrators involved in disciplinary outcome decisions will be asked to take a short online test and survey asking about their years of experience and racial identification and measuring their unintentional (or implicit) bias based on race. Finally, you will be asked to share your school district's percentage of students of color, percentage of students eligible for free or reduced lunch, and the total number of students in the district. All information will remain confidential for administrators and schools and anonymous for students (student names should never be shared with the researcher).

Risks and *Benefits* of being in the study
Possible risks:
First, an unintentional change in student discipline severity may occur due to an increased awareness of bias. This is unlikely as bias test results will not be shared during the study, but only if administrators choose to retake the test as provided with debriefing materials. These materials will help participants understand individual bias test results and how to lessen the effects of any biases noted. Second, administrators taking the bias test might experience some discomfort or guilt because of exposure to results. This again is both optional and lessened by the debriefing materials. Although this risk is more prevalent, it is short-term and often more educational in nature than disconforting. This is a one-time risk during the study and may be experienced again only if the participant decided to retake the test after the study to learn of his or her results.

The benefits to participation are:
An increased awareness of personal bias through debriefing measures of bias provided to participants and access to an equity audit of shared district data. This data will demonstrate any inequities in discipline by traditionally underserved populations (students of color or those students eligible for free or reduced lunch). With this data, schools can monitor discipline decisions to develop the most equitable environment for student to learn as possible.

Compensation
You will receive payment:
In the form of a district equity audit using the provided discipline data. Administrator bias test participation is not necessary to receive this compensation. All districts and schools sharing any data will receive this anonymous analysis at the conclusion of the study. This report includes percentages and graphs to help schools know if discipline is being equitably administered in the district and if behavioral infractions are equitably occurring

Confidentiality

The records of this study will be kept private. In any sort of report we might publish, we will not include any information that will make it possible to identify a subject. Research records will be stored securely and only researchers will have access to the records.

Voluntary Nature of the Study

Participation in this study is voluntary.
Your decision whether or not to participate will not affect your current or future relations with the Lehigh University. If you decide to participate, you are free to not answer any question or withdraw at any time without affecting those relationships.

Contacts and Questions

The researchers conducting this study are:
Gina L. Gullo and Dr. Floyd Beachum. You may ask any questions you have now. If you have questions later, **you are encouraged** to contact them at Iacocca Hall Room A-210, 610-758-5955, FDB209@Lehigh.edu.

Questions or Concerns:
If you have any questions or concerns regarding this study and would like to talk to someone other than the researcher(s), **you are encouraged** to contact Naomi Coll, Lehigh University's Manager of Research Integrity, at (610) 758-2985 (email: nac314@lehigh.edu). All reports or correspondence will be kept confidential.

You will be given a copy of this information to keep for your records.

Statement of Consent

I have read the above information. I have had the opportunity to ask questions and have my questions answered. I consent to participate in the study.

Signature: _____ Date: _____

Signature of Investigator: _____ Date: _____

181

APPENDIX D

Site Permission Form

Site Permission Letter

This agreement is limited to the following specific protocol:

Name of Research Project: Implicit Bias in School Disciplinary Decisions
Protocol Number: 851380-1
Name of Principal Investigator: Floyd Beachum, Ed.D.
Name of Researcher: Gina Gullo, M.Ed.

Research Scope: To better understand how unintentional bias might be involved in decision making for Pennsylvania principals and assistant principals when making decisions about behavior consequences for discipline infractions.

Activities to be conducted at/by Site: School discipline data routinely collected by the site will by blinded in order to prevent student identification and shared electronically with the researcher. This data will include a student identifier (not name), student grade, student race/ethnicity, student behavioral infraction, student discipline consequence, and administrator determining discipline consequence. School administrators involved in discipline will complete an implicit association test electronically at their leisure (possibly on site, but available anywhere internet access is available). They will also complete two questions regarding experience and race/ethnicity identification.

Other: Participating schools/districts will receive a complementary Equity Audit Report generated from the shared data. This report will not name any particular administrators. The schools will also be given access to an online tool to measure implicit and explicit bias and offer suggestions for working through both kinds of biases towards a more equitable school environment.

School/District Name:_____

Print Full Name: _____

Institutional Title: _____

By signing below, I give permission for the above research to be conducted with my site. I also certify with my signature below that I have the appropriate authority to provide this permission for the school or school district named above.

Signature: _____

Date:_____

4425 Charles Street
Easton, PA 18045
@ Lehigh: Iacocca Hall A-208

201.618.3075
Gina.Ciani@gmail.com
http://wordpress.lehigh.edu/glc211

EDUCATION

Lehigh University, College of Education   Expected May 2017
Ed.D., Educational Leadership
Dissertation: Implicit Bias in School Disciplinary Decisions (Defend April 20, 2017)

Lehigh University, College of Education   May 2014
M.Ed., Special Education

Rutgers University   May 2008
Complete Masters Coursework, Behavioral Neuroscience

The College of New Jersey   December 2005
B.A., Biopsychology
Minors: Deaf Education, Computer Science

PUBLICATIONS

Gullo, G. L. (in press). The risk of discipline for youth in need. In F.D. Beachum & F. E.
Obiakor (Eds.). *Improving Educational Outcomes of Vulnerable Children:
Starting from the Bottom.* San Diego, CA: Plural Publishing Inc.

Beachum, F. D., Obiakor, F. E., & Gullo, G. (in press). Understanding vulnerable
populations: From perceptions to realities. In F.D. Beachum & F. E. Obiakor
(Eds.). *Improving Educational Outcomes of Vulnerable Children: Starting from
the Bottom.* San Diego, CA: Plural Publishing Inc.

Beachum, F. D., & **Gullo, G. L.** (2014). Promoting socio-cultural justice for culturally
and linguistically diverse secondary school students. In Zhang, C., McCray, C. R.,
& Cho, S. (Eds.). *Effective education for all: Implementing positive behavior
support in early childhood through high school*. New York, NY: Peter Lang
Publishing.

WORKING PUBLICATIONS

Gullo, G. L. (under review). Using data for school change: The discipline equity audit
and school climate survey. *Journal of Cases in Educational Leadership.*

Gullo, G. L., & Sperandio, J. (pending revisions). Gender and the superintendency: The role of career pathways. *Educational Administrator Quarterly.*

Gullo, G. L. Capatosto, K., & Staats, C. (in proposal). *Implicit Bias in Schools: Bridging Research to Practice.* New York, NY: Routledge.

K-12 TEACHING AND CONSULTING

Discipline Equity Audits (Pro-Bono) Various PA School Districts	2016-2017
School Climate Data Analysis Allentown School District, PA	2015
Equity Audit Data Analysis Allentown and Bethlehem Area School Districts, PA	2015-2016
Substitute Teaching (Source4Teachers, Substitute Teaching Service) K-12 students, special and general education Urban, Suburban, and Rural Pennsylvanian schools	2014-2016
Math and Reading Tutor (Pro-Bono; Bethlehem Area School District, PA) Worked with low-income students on math and reading interventions	2013-2015
Science Afterschool Program Instructor (Mad Science) PK-5 students, 1-3 hour programs	2003-2006; 2011-2014
Student Teaching (Bethlehem Area School District) Grade 4 general education; K-3 Autistic Support	2014
Summer Preschool Teacher (Canaan Presbyterian Preschool) English immersion preschool for Korean students	2002-2004

HIGHER EDUCATION TEACHING AND LEADERSHIP

Graduate Student Senate Lehigh University, 2016-17 Historian, 2012-13 Communications Officer	2011-2017
Graduate Student Fathers and Mothers Club Lehigh University, Founder and President	2016-2017

Lehigh University Committees
  Graduate Diversity and Inclusion Committee, Chair 2016-17
  Graduate Student Senate Constitutional Reform Committee, Chair 2016-17
  Faculty Committee for Student Life, Graduate Student Rep 2012-13
	2012-2017

Co-Instructor (Introduction to Statistics) Lehigh University, (Under Supervision of Dr. Qiong (Joan) Fu)	2016

Co-Instructor (Qualitative Methods)	2016
Lehigh University, (Under Supervision of Dr. Jill Sperandio)	

Co-Instructor (Qualitative Methods)   2016
  Lehigh University, (Under Supervision of Dr. Jill Sperandio)

Teaching Assistant (Leading Inclusive Learning Systems)   Annually 2011-2016
  Lehigh University, Instructor: Dr. Floyd Beachum

Teaching Assistant (Diversity and Multicultural Perspectives)   Annually 2012-2016
  Lehigh University, Instructor: Dr. Floyd Beachum

Teaching Assistant (Instructional Leadership)   Annually 2012-2016
  Lehigh University, Instructor: Dr. Floyd Beachum

Laboratory Instructor (Sensation and Perception)   2007, 2008
  Rutgers University

Laboratory Instructor (Cognitive Psychology)   2008
  Rutgers University

CONFERENCE PRESENTATIONS

Demonstration, May 2016: "Implicit Bias in Decision Making"
- Lehigh University Graduate Experience Expo

Roundtable, November 2016: "Behavioral Decision-Makers' Perceptions of Exclusionary Discipline Decisions and the Influence of Implicit Bias"
- University Council for Educational Administrators (UCEA) Graduate Student Summit

Paper, November 2016: "Breaking Gender Walls through Pathways to the Superintendency"
- University Council for Educational Administrators (UCEA) Conference

Poster, October 2016: "Principals Acting as Instructional Leaders despite External Constraints"
- The College of New Jersey Interdisciplinary Research Forum

INVITED PRESENTATIONS

Presenter: Lehigh University College of Education Diversity and Inclusion Lectures (2017)
- Implicit Bias: Measurement and Mitigation

Guest Lecturer: Diversity and Multicultural Perspectives (2017)

- The Implicit Associations Test
- Using Data for School Reform

Guest Lecturer: Leading Inclusive Learning Systems (2016)
- The Implicit Associations Test

Guest Lecturer: Doctoral Seminar: The Literature Review (2016)
- Analyzing Qualitative and Quantitative Articles

- Organizing Literature and Time Management

Presenter: Transition to Graduate School Series (2015, 2016)
- Imposter Syndrome and Time Management

Guest Lecturer: Advanced Multimedia Programming and Resource Development (2013)
- Using Voice Overs in Educational Multimedia

AWARDS, CERTIFICATES, & GRANTS

VISIONS Language/Framework Training (February 2017)

Lehigh University Ally (LGTBQ+ Support; 2016-2017)

Lehigh University College of Education Diversity Travel Grant ($300; 2016)

Lehigh University Graduate Student Travel Grant ($150; 2016)

Lehigh University Dean's Endowed Student Travel Scholarship ($400; 2016)

Institutional Research Board – Human Subjects Certification (2015)

Lehigh University College of Education Graduate Life and Leadership Award ($200; 2014)

Bergen County Community Service Award (2012)

College Teaching Training Level I and II (2012)

Microsoft Office 2010 Suite – Advanced Certification (2011)

The College of New Jersey Student Involvement Award ($75; 2004)

ADDITIONAL EDUCATION-RELATED EXPERIENCE

Graduate Experience Expo Planning Committee   2017
Lehigh University

Vice Provost Hiring Committee   2017
Lehigh University, Student Consultant

Legal Data Analyst   2012-2017
Independent/Lehigh University for Dr. Perry Zirkel

Certificate for Inclusive Excellence Committee   2016-2017
Lehigh University

Research Assistant to Dr. Floyd Beachum   2012-2017
Lehigh University, Educational Leadership

Educational Leadership Faculty Hiring Committee   2016
Lehigh University, Student Representative

Vice President for Equity and Inclusion Hiring Committee   2016
Lehigh University, Student Consultant

Gender and the Superintendency: Perceptions of Career Pathways   2014-2016
Interview follow-up to previous study; Qualitative analyses

RAMP-UP (Reading Achievement Multi-Component Program)   2015
Conducted Woodcock-Johnson tests on students

CARS (Center for Adolescences Research in Schools)   2011-2015
Collected behavioral observation data on "at-risk" students

Lehigh Community Support   2011-2012
Cared for seniors with special needs in group home

Bring Your Child to Work Day Coordinator (Huntingdon Life Sciences)   2009-2011
Planned, scheduled, and facilitated programming for special event

College Teaching Excellence Course Completion   2007
Rutgers University, NJ

COMMUNITY INVOLVEMENT

Courageous Conversations Series   2017
- Planned, organized, and led a series of three events at Lehigh University
- Focused on issues of conflict surrounding Islam, race, and LGTBQ+ communities

Lehigh University Graduate Student Constitution Committee   2016-2017
- Founder 2016
- Chair 2016-2017
- Task Force to revitalize Graduate Student Senate constitution

Lehigh College of Education Dissertation Coffee Hour   2016-2017
- Founder, Coordinator

Lehigh University Fathers and Mothers Club   2016-2017
- Founder, President 2016-2017
- Member 2016-2017

Lehigh University Graduate Student Diversity and Inclusion Committee   2015-2017
- Co-Chair 2016-2017
- Member 2015-2017

Christ United Methodist Church of Easton, PA   2013-2017
- Active Member, Vacation Bible School Teacher

Lehigh University Graduate Student Senate   2011-2017
- Historian 2016-2017
- Representative 2011-2017
- Communications Officer 2012-2013

Easton Farmer's Market   2009-2017
- Local Community Service, Volunteer
- Help to judge and coordinate culinary competitions
- Help organize and manage children's programming
- Strategize for public involvement

Easton Garlic Fest   2016
- Helped with culinary organization and planning
- Assisted and judged children's cooking competitions

PA Bacon Fest   2004, 2016
- Participated in cooking competitions
- Assisted with press interactions and advertising

Musikfest   2014-2016
- Local Community Service, Volunteer
- Helped exchange money for tickets

Lehigh University V-Day Initiative   2011-2015
- Helped manage and perform "Vagina Monologues"
- Mentored undergraduates on being a woman in academia
- Raised funds for local women/children abuse shelters

Lehigh University Faculty Committee for Student Life   2012-2013
- Graduate Student Representative 2012-2013
- Managed and awarded grant money