

University of Northern Colorado

UNCOpen

---

Dissertations

Student Work

---

8-2024

## Combating the Racial Discipline Gap: Does Diversity Training Mitigate Implicit Bias in the Discipline Decision- Making Process?

Bailey R. Reshel

*University of Northern Colorado*

Follow this and additional works at: <https://digscholarship.unco.edu/dissertations>

---

### Recommended Citation

Reshel, Bailey R., "Combating the Racial Discipline Gap: Does Diversity Training Mitigate Implicit Bias in the Discipline Decision- Making Process?" (2024). *Dissertations*. 1105.

<https://digscholarship.unco.edu/dissertations/1105>

This Dissertation is brought to you for free and open access by the Student Work at UNCOpen. It has been accepted for inclusion in Dissertations by an authorized administrator of UNCOpen. For more information, please contact [Nicole.Webber@unco.edu](mailto:Nicole.Webber@unco.edu).

© 2024

BAILEY R. RESHEL

ALL RIGHTS RESERVED

UNIVERSITY OF NORTHERN COLORADO

Greeley, Colorado

The Graduate School

COMBATING THE RACIAL DISCIPLINE GAP: DOES  
DIVERSITY TRAINING MITIGATE IMPLICIT  
BIAS IN THE DISCIPLINE DECISION-  
MAKING PROCESS?

A Dissertation Submitted in Partial Fulfillment  
of the Requirements for the Degree of  
Doctor of Philosophy

Bailey R. Reshel

College of Education and Behavioral Sciences  
Department of Applied Psychology and Counselor Education  
Program of School Psychology

August 2024

This Dissertation by: Bailey R. Reshel

Entitled: *Combating the Racial Discipline Gap: Does Diversity Training Mitigate Implicit Bias in the Discipline Decision-Making Process?*

has been approved as meeting the requirement for the Degree of Doctor of Philosophy, in College of Education and Behavioral Sciences, Department of Applied Psychology and Counselor Education, in the Program of School Psychology

Accepted by Doctoral Committee

---

Robyn Hess PhD, Research Advisor

---

Michelle Athanasiou PhD, Committee Member

---

Kristin Myers PhD, Committee Member

---

Amie Cieminski PhD, Faculty Representative

Date of Dissertation Defense \_\_\_\_\_

Accepted by the Graduate School

---

Jeri-Anne Lyons, Ph.D.  
Dean of the Graduate School  
Associate Vice President for Research

## ABSTRACT

Reshel, Bailey. *Combating the racial discipline gap: Does diversity training mitigate implicit bias in the discipline decision-making process?* Published Doctor of Philosophy dissertation, University of Northern Colorado, 2024.

The racial discipline gap between Black and White students has been pervasive for decades. To date, limited research has been conducted on the possible role that principals' implicit racial biases may play in this gap. As such, the purpose of this study was to explore how high school principal's implicit racial biases may influence the severity of consequences given to Black students during subjective discipline situations. A non-experimental quantitative approach was used to explore the possible impact of high school principals' pro-White implicit racial biases during subjective disciplinary decisions with Black students. It additionally examined whether the number of hours of diversity training a principal has engaged in might be a mitigating factor for pro-White biases during discipline with Black students. To conduct this study, 166 high school principals completed an online survey during which they responded to three disciplinary vignettes with either a common White name, a common Black name, or a name that is common across races. They were then asked to estimate the hours of diversity training they had completed over the past three years followed by the completion of a name-based implicit association test (IAT). The latter information was used to answer the following three research questions:

- Q1     After accounting for hours of diversity training, is principal implicit bias a significant predictor of discipline severity on Black student vignettes?

Q2 Do differences exist between the severity of consequences assigned to Black and White student vignettes?

Q3 Do differences exist between the severity of consequences assigned to Black and White student vignettes accounting for hours of diversity training?

Results indicated that implicit racial biases were not a predictor of the discipline severity of the hypothetical Black student. Additionally, the amount of diversity training a principal reported was not predictive of lesser consequences for the hypothetical Black student either. After analysis of the findings, it was hypothesized that name-based measures (i.e., vignettes and IATs) may not be sufficiently activating enough to elicit participant responses that are reflective of real-life subjective disciplinary scenarios.

## DEDICATIONS

I am so grateful for my family for their unconditional love and support through this process. Mom and dad--thank you for giving me a landing pad during all the adventures and challenges I have charged head on and for your encouragement to pursue each and every one. This accomplishment is just as much yours as it is mine. Beausephine--thank you for every laugh and heart to heart I needed to persevere. I would be completely lost in this life without you. Love-love. Luka--thank you for the joy, companionship, and endless laughs you have gifted me. I could not have asked for a more patient and gentle soul to have by my side through this wild ride. I love you, peanut.

## TABLE OF CONTENTS

CHAPTER		
I.	INTRODUCTION .....	1
	School Discipline Policies .....	2
	Theoretical Framework .....	3
	Purpose of the Study .....	5
	Research Questions .....	7
	Delimitations .....	8
	Definition of Terms .....	9
	Conclusion .....	12
II.	REVIEW OF LITERATURE .....	13
	The School-to-Prison Pipeline .....	13
	Disproportionality and the School-to-Prison Pipeline .....	16
	Contributing Factors .....	19
	School Policies .....	21
	Bias....	24
	Existing Research on Principal Implicit Bias and Discipline .....	26
	Bias and Subjectivity .....	30
	Principal Preparation Programs and Professional Standards .....	31
	Principal Preparation Programs .....	33
	Continuing Education and Practicing Principals .....	36
	Building Culturally Responsive Practitioners .....	38
	Summary .....	41
III.	METHODOLOGY .....	42
	Instrumentation .....	42
	Demographic Information .....	43
	Vignettes .....	44
	Implicit Association Test .....	45



## CHAPTER

### III. continued

	Procedure .....	51
	Research Questions .....	51
	Data Analysis .....	52
	Summary .....	53
IV.	RESULTS .....	54
	Description of Participants.....	54
	Implicit Racial Bias.....	58
	Discipline Severity.....	60
	Research Question 1 .....	62
	Research Question 2 .....	65
	Research Question 3 .....	68
	Summary .....	75
V.	DISCUSSION .....	76
	Overview of Results.....	76
	Implications.....	79
	Limitations .....	81
	Future Research .....	82
	Conclusion .....	85
	REFERENCES .....	86
	APPENDIX	
A.	Content Permissions.....	104
B.	Demographic Survey Questions .....	106
C.	Discipline Vignettes.....	109
D.	Race Implicit Association Task .....	111
E.	Institutional Review Board Approval Letter.....	117

## LIST OF TABLES

Table		
1.	Subjective Vignettes and Consequence Options.....	45
2.	Implicit Association Test Stimuli .....	47
3.	Standard Seven-Block Procedure .....	49
4.	D Score Quantitative Categories and Qualitative Descriptors .....	51
5.	Summary of Participant Demographic Information .....	57
6.	Summary of Participant Implicit Racial Bias D Scores.....	59
7.	Discipline Consequence Frequencies Assigned by Vignette.....	61
8.	Correlations for Discipline Rating, Vignette Number, Hours of Diversity Training, and D Score.....	64
9.	Hierarchical Multiple Regression Exploring the Discipline Severity of the Name Dariu .....	s 66
10.	Independent Samples <i>t</i> -Test Comparing the Discipline Severity Between Darius and Brian .....	68
11.	Hierarchical Multiple Regression Predicting Differences in Discipline Severity of Darius and Brian When Accounting for Hours of Diversity Training.....	70

## LIST OF FIGURES

Figure

1. The Paradox for the Principals Understanding Social Justice ..... 35

## CHAPTER I

### INTRODUCTION

For decades, Black students have consistently been the most overrepresented population in exclusionary school discipline (Carter et al., 2017; DeMatthews, 2016b; Diem & Welton, 2020; Fenning et al., 2008; Gregory et al., 2010); however, there has been no evidence to suggest that Black students have engaged in problematic behavior at higher rates than other student populations (Diem & Welton, 2020; Sander et al., 2011). Taken together, this is concerning because exclusionary school discipline practices often increase the risk of juvenile incarceration, a concept referred to as the school-to-prison pipeline (STPP), which is more common for Black youth (Diem & Welton, 2020). Additionally concerning is that, although the national rate of juvenile incarceration has declined substantially for all races in the past 20 years (Hockenberry, 2020), the disparity between Black youth and those of other races has remained stable (Annie E. Casey Foundation, 2021; Diem & Welton, 2020; Redfield & Nance, 2016; Rovner, 2021; Ruiz, 2017; The Sentencing Project, 2021). These persistent trends have contributed to the hypothesis that when students are pushed out of their schools through exclusionary disciplinary practices, they are more likely to become involved in negative behaviors due to the lack of structure and support they would otherwise receive if they remained in school; this exclusion and its subsequent consequences then increases the likelihood of incarceration. To better understand the reasons behind such outcomes, this study seeks to understand principals' implicit racial bias, professional training focused on diversity, and disciplinary decision-making with students from diverse racial backgrounds and how each of these may assist in better understanding the STPP.

### **School Discipline Policies**

Nearly 95% of school suspensions in the United States have been for “nonviolent, minor disruptions” (Llorente, 2014, p. 15), which has been a phenomenon attributable to zero tolerance policies (Diem & Welton, 2020; Ruiz, 2017; Skiba, 2014). Such policies were initially implemented in the 1990s to address egregious violations in schools, especially school violence. They were widely adopted by districts across the country and applied to any number of offenses, not just egregious ones, which eventually led to strict school disciplinary practices for misbehavior of any degree becoming the norm. Unfortunately, because there are associated financial incentives for schools to maintain these practices (Welch & Payne, 2018), zero tolerance policies are still commonplace 30 years after their initial enactment. They continue to be detrimental to the student body at large but have tended to be especially harmful for Black students who are referred to the office and suspended twice as often as White students (Girvan et al., 2016; Owens & McLanahan, 2020; Van den Bergh et al., 2010). Often, such referrals are for subjective, non-violent offenses such as perceived dress code violations, tardiness, truancy, defiance, or arguing, to name a few. The discrepancy in disciplinary referrals between Black and White students is referred to as the racial discipline gap and has continued to persist even after accounting for and intervening at the student and teacher level (DeMatthews et al., 2017). Thus, there is a need to explore other aspects of the school system that could be contributing to the racial discipline gap, such as principals’ role in the disciplinary process.

In school systems, principals hold notable positions of power. They decide which policies to prioritize and reinforce among staff, are pivotal in shaping the climate of their schools, pick the in-service trainings for staff, and, most important to this study, are pivotal to disciplinary decision-making (National Policy Board for Educational Administration, 2015). Principals’

personal attitudes and beliefs about these different responsibilities likely influence how they carry out each in their schools. Regarding disciplinary decision-making, during situations where there may not be a lot of information to work with, principals may feel pressured to act in favor of efficiency or in ways that support their staff in lieu of more thoroughly examining a disciplinary concern. When either of these occur, principals may use detrimental (albeit unintentional) mental shortcuts to navigate disciplinary decisions, an example of such being unconscious or implicit biases. For illustration, individuals may unconsciously allow the negative attitudes against other groups (e.g., females, Black individuals, lawyers, police officers) to guide their thinking, even if they are not consciously aware of these negative attitudes. The phenomenon of having negative attitudes about a group of people that are out of conscious awareness is rooted in a concept called implicit bias (Carter et al., 2017). Because Black individuals are a group with a long history of systemic marginalization in the United States, the resulting implicit biases of this history may help explain why Black students continue to receive discipline referrals twice as often as White students (Brown, 2018). For this reason, it is worth examining how implicit biases might influence disciplinary decision making amongst principals.

### **Theoretical Framework**

Implicit biases are subconscious mental shortcuts that influence how people think about and respond to their environments (Bursell & Olsson, 2021; Carter et al., 2017; Ruhl, 2021). They are learned thought patterns that increase cognitive efficiency due to their automaticity. Although not inherently bad, they can be harmful when their automaticity leads to overapplication. An example of this is stereotyping, which is rooted in the concept of “otherness”, an especially strong adaptive mental shortcut that allows humans to quickly categorize people into groups (e.g., race, ethnicity, education level, socioeconomic status,

gender, age; National Center for Cultural Competence, n.d.) in order to efficiently organize their social worlds. The danger of such thought processes is that they can be activated without sufficient thought or reflection to the specific attributes of *individuals* and lead to inaccurate overgeneralizations of a broader group (e.g., girls are more social; Asian students are smart, Black men are dangerous; Ruhl, 2021).

As implicit biases and stereotyping pertain to this study, there is a longstanding, deeply engrained negative stereotype of Black individuals in the United States. Combined with the otherness shortcuts that all humans use, there has been high potential for many individuals who are not Black to hold negative implicit biases about this demographic. As it pertains to schools, the racial makeup of staff may be of particular relevance. Approximately 80% of all principals and teachers in K-12 schools in the United States have been White and were thought to have the same levels of implicit racial biases as the general population (McIntosh et al., 2021; Minkos et al., 2017; National Center for Education Statistics [NCES], 2019, 2020a, 2020b). For minoritized students, the lack of diversity in school staff could be problematic because research has shown that simply being in the presence of someone of a race other than one's own is enough to elicit harmful stereotyping thought processes--even if they do not align with one's consciously stated beliefs (Brown, 2018; Greenwald & Krieger, 2006; Gullo et al., 2018). A compounding factor is the interplay of implicit biases and environmental stressors. Secondary schools tend to be fast-paced environments where efficiency is often prioritized and encouraged. In such conditions, shortcuts are much more likely to drive decision-making in order to increase efficiency. Furthermore, in high stress and efficiency driven environments, principals are prone to high cognitive load (Fenning et al., 2008; Gullo et al., 2018). When cognitive load is high, the likelihood of cognitive shortcuts (i.e., stereotyping) being activated increases in order to promote

mental efficiency. As it pertains to principals, even the most well-intentioned individuals who endorse beliefs in fair discipline practices are susceptible to acting in unconsciously harmful ways due to the nature of the environments they work in (Gullo et al., 2018).

Lastly, and most important to this study, is that mental shortcuts are more likely to be activated when a situation contains a higher degree of subjectivity, meaning it contains interpretive ambiguity that is more susceptible to the influence of personal attitudes and beliefs, including implicit ones (Girvan et al., 2016). In schools, subjective infractions (e.g., disruption, defiance, misconduct, disrespect, insubordination) of the same type are more likely to have differing consequences due to the variation across the individuals making disciplinary decisions. Furthermore, when information is lacking or a situation is perceived as ambiguous, mental shortcuts (i.e., stereotypes, implicit biases) help fill informational gaps during subjective disciplinary situations (Greenwald & Krieger, 2006). Given the interplay between environmental demands, deeply engrained societal Black stereotypes (Gullo et al., 2018), the racial homogeneity of secondary principals, and the automaticity of mental short cuts, there is high potential that implicit biases influence disciplinary decisions.

### **Purpose of the Study**

The racial discipline gap has persisted for over 30 years. Despite researchers' attempts at understanding this phenomenon, the gap is still difficult to explain even after studying student, teacher, and community level factors (Carter et al., 2017; DeMatthews, 2016b; Diem & Welton, 2020; Fenning et al., 2008; Gregory et al., 2010). Thus far, student's race and school level factors (i.e., school size, population demographics) have best predicted the likelihood of student suspension (Gullo, 2017); however, why these factors have been the best explanation thus far is still not well understood. A reasonable next step is to examine the potential role of principals in



racial discrepancies considering they are the last stop in the disciplinary process. To date, the majority of studies examining principal disciplinary decision making have been qualitative, not representative of the general population of principals, and not representative of school districts at large (Jarvis & Okonofua, 2020). In the few quantitative studies that exist, principal's pro-White preferences and racial implicit biases have accounted for some of the differences in disciplinary harshness between White and minoritized students (Gullo, 2017; Gullo & Beachum, 2020a, 2020b; Jarvis & Okonofua, 2020). Such results help strengthen the hypothesis that implicit biases may be a predicting factor in suspension, but further exploration is needed. Additionally, of the few qualitative studies that currently exist, there are none, to this researcher's knowledge, that simultaneously explore the existence and/or impact of implicit racial bias in disciplinary decisions *and* potential mitigating factors. To further inform this body of research, the current study examined whether principals assign harsher disciplinary consequences to Black students because of pro-White implicit racial biases. The researcher did so while also exploring how formal training experiences might lessen racial discipline discrepancies with the hope of informing a possible intervention option for combatting the STPP.

### **Need for the Study**

The need for this study was two-fold. First, with the implementation of the Every Student Succeeds Act (ESSA; 2015), the U.S. Department of Education called on the institution of public education to strive for the advancement of equity. As it pertains to the racial discipline gap, there is still a long way to go in achieving equity. Thus, better understanding how the interplay of principal implicit biases and the kinds of training they receive impacts the racial discipline gap is needed, especially because of their gatekeeping role in the disciplinary process. Second, because there will always be a relational component to disciplinary processes, there is utility in better

understanding the impact of cultural training on decision making. By better understanding what personal factors (i.e., racial implicit bias) of principals could harm disciplinary processes *and* how to treat possible relational barriers, it may be possible to inform targeted interventions for mitigating the racial discipline gap and the STPP. Exploring what fosters culturally competent leaders as principal preparation programs and continuing education options for practicing principals evolve is a potential way of exploring such interventions.

### **Research Questions**

This study used a quantitative approach to understanding the potential role of principal racial implicit bias in disciplinary decision-making. It was intended to add to pre-existing research by exploring whether diversity training could be a viable option for reducing the effect of negative racial implicit bias. Although previous research studies have explored similar questions, the qualitative methodologies (e.g., sample study) that were used provided depth but not generalizable findings. This study was designed to survey a large number of secondary principals in the Midwest in order to understand broader trends that may inform school practice and training approaches that contribute to more equitable decision-making.

The following research questions were used:

- Q1     After accounting for hours of diversity training, is principal implicit bias a significant predictor of discipline severity on Black student vignettes?
- Q2     Do differences exist between the severity of consequences assigned to Black and White student vignettes?
- Q3     Do differences exist between the severity of consequences assigned to Black and White student vignettes accounting for hours of diversity training?

Two statistical methods were used to answer the three research questions. These included hierarchical multiple regression to examine whether principals' racial implicit biases would predict the discipline severity of Black students, an independent-samples *t*-test to analyze the

difference between the severity of consequences assigned on Black and White student vignettes, and additional hierarchical multiple regressions to see if there were differences in the severity of consequences on Black and White student vignettes after accounting for hours of diversity training.

### **Delimitations**

The STPP is a very broad, multifaceted concept that is too complex to explain or explore with just one research study. Student, family, school, community, and state and federal policies all contribute to its existence. Because a thorough examination of each of these factors on the disciplinary process was beyond the scope of this study, only principals were included as participants as a way to provide a much-needed, highly focused understanding of one component of the STPP.

A second consideration pertains to the use of disciplinary vignettes. A core assumption of this study was that implicit biases can differ from explicit biases (Greenwald & Krieger, 2006). Thus, questionnaires explicitly asking principals about their racial beliefs or attitudes as they pertain to disciplinary decision making was moot for two reasons. First, not only might participants be unaware of their unconscious biases, but second, they may also respond in socially desirable ways if they were aware of them. In an attempt to account for these possibilities, vignettes were intended to be short enough to create reasonable ambiguity that might activate implicit biases that would also be active during real-life decision-making scenarios. However, there was the possibility that participants would recognize the strategically chosen names for vignettes and respond in socially desirable ways despite the researcher's attempts to disguise the overarching theme of the study. There was also the potential that the vignettes themselves would not be sufficient to elicit the implicit biases that may be present

during real-life interactions. For these reasons, it was necessary to strategically describe the study in a way that guided participants to believe this study was about the discipline making process as a whole rather than an exploration of how student race may affect decision making.

A final consideration was the use of Implicit Association Test (IAT) as the method for measuring implicit racial bias. There have been mixed opinions among social science researchers about the validity and usefulness of IATs (Sleek, 2018). Primary concerns have included the test's validity and whether they are superior to other more explicit self-report measures (Oswald et al., 2015; Schimmack, 2021a, 2021b). However, there is a body of empirical research supporting their validity and utility when looking at *specific subcategories* of IATs (e.g., race, gender, children, politics, risk taking, obesity, etc.; Kurdi, Carrol, & Banaji, 2021; Kurdi, Ratliff, & Cunningham, 2021; Tosi et al., 2020). Therefore, the use of a race IAT was deemed appropriate for this study. Additionally, the use of an IAT was considered reasonable due to its results being interpreted within the context of other sources of data.

### **Definition of Terms**

*Cultural responsiveness.* Cultural responsiveness is an ongoing process that entails leading school systems and interacting with members of a school system in a way that encompasses awareness of how one's varying identities (e.g., race, ethnicity, gender, political affiliation, religion, economic status) may differ from those of others. It also means being cognizant of how within- and between-group differences can influence decisions and interactions. Additionally, cultural responsiveness means having respect for all students and families as well as displaying an ongoing effort to understand individual and group differences with the intent to best meet student needs (Bal et al., 2018;

Metropolitan Council of Educational Administration Programs Committee on Cultural Responsiveness, 2017).

*Diversity training.* Any training that aims to foster cultural responsiveness by “[facilitating] positive group interaction, [reducing] prejudice and discrimination, and generally [teaching] dissimilar others how to work together effectively” (Lindsey et al., 2014, p. 606). Diversity trainings are designed to foster self-reflection of one’s racial and cultural identities, challenge preexisting knowledge about others’ experiences and identities, and encourage the dismantling of systems of oppression through positive interpersonal interactions (Bezrukova et al., 2012; Jackson et al., 2014).

*Exclusionary discipline.* Any disciplinary practice that removes students from their usual education setting (Diem & Welton, 2020). This term includes any type of discipline that removes a student from their typical learning environment (e.g., in school suspension, out of school suspension, expulsion).

*Explicit bias.* Explicit biases are feelings, attitudes, and behaviors that are processed at a conscious level and known to an individual (National Center for Cultural Competence, n.d.).

*Expulsion.* Permanent removal of a student from a school or removal for an extended period of time that includes return contingencies.

*Heuristics.* Mental shortcuts used to improve cognitive efficiency during decision making (Fiske & Taylor, 2008).

*Implicit Association Test (IAT).* An interactive task used to quantify possible implicit racial biases. For this study, a name-based IAT was used that included stereotypical Black and White names along with pleasant and unpleasant words.

*Implicit bias.* Implicit biases are subconscious, pervasive, automatic mental processes that influence how people think about and respond to their environments (Carter et al., 2017; Ruhl, 2021).

*National Policy Board for Educational Administration.* A national alliance of membership organizations dedicated to advancing the field of school leadership.

*Office discipline referral.* The referral of a student to the principal's office that is initiated by teachers or school staff for student disciplinary concerns. Typically, these referrals are for repeated infractions or more serious rule violations that are beyond what teachers or other school staff members can address.

*Principal.* Within the context of this study, a principal was an all encompassing term that consisted of school leaders that are responsible for making student disciplinary decisions. For simplicity, this included principals in addition to vice principals, assistant principals, and deans.

*Racial discipline gap.* A phrase used to describe the discrepancy in disciplinary consequences given to minoritized students in comparison to White students (Diem & Welton, 2020; Fenning & Rose, 2007; Gregory et al., 2010; Owens & McLanahan, 2020).

*School-to-prison pipeline.* A concept that captures a student's entry into the justice system that either starts with or is exacerbated by exclusionary discipline practices.

*Social justice.* The concept of actionably advocating for equity. In the context of schools, social justice refers to acting in ways that foster equal access, respect, and fairness for all students (Shriberg et al., 2013).

*Suspension.* Temporary removal of a student from their typical educational setting.

*Zero tolerance policies.* Policies enacted in response to the Gun Free School Act of 1994 to combat concerns about school violence (Ruiz, 2017). Such policies rely heavily on exclusionary discipline as a method of enforcement (Diem & Welton, 2020).

### **Conclusion**

The overrepresentation of Black students in exclusionary discipline and the justice system is an ongoing concern. There are many factors that have contributed to this disturbing trend, and a deeper examination of school processes may add to the existing body of research on what is already known about the STPP. More specifically, one research avenue that has been minimally explored is the potential interplay between principals' diversity training and racial implicit biases and how it may influence disciplinary decision making with Black students. As the disciplinary gatekeepers, it is important to better understand how principals' decisions contribute to the continued racial discipline gap. By exploring the prevalence of implicit biases in principals, and whether targeted training offsets harmful implicit biases, there was opportunity to inform potential interventions for mitigating the STPP.

## CHAPTER II

### REVIEW OF LITERATURE

In the last 20 years, the overall national juvenile incarceration totals in the United States have consistently declined from 105,055 in 1997 to 43,580 in 2017 (Hockenberry, 2020). Yet, this success has been overshadowed by the fact that Black youth are still significantly overrepresented in the justice system (Diem & Welton, 2020; Redfield & Nance, 2016; Rovner, 2021; Ruiz, 2017; The Sentencing Project, 2021). Despite comprising only 15% of the overall youth population, Black individuals make up 41% of incarcerated juveniles (Rovner, 2021). Many factors have contributed to this discrepancy, but exclusionary school discipline policies in particular have come under scrutiny. Entry into the justice system that either starts with or is exacerbated by exclusionary school discipline is referred to as the STPP and Black students experience such practices at disproportionate rates.

#### **The School-to-Prison Pipeline**

As defined by the National Association of School Psychologists (NASP, 2019), the STPP refers to the concept of schools funneling students into the justice system via exclusionary discipline practices. Diem and Welton (2020) defined exclusionary discipline as “any form of disciplinary practice that results in removing or excluding a student from their typical educational setting” (p. 118). Examples have included in-school suspension, out of school suspension, and expulsion. Aside from office discipline referrals (ODRs), suspension is the most common disciplinary action taken against students (Fenning & Rose, 2007). Once removed from the classroom due to a suspension or expulsion, students are susceptible to falling behind in their



classes, spending hours each day without supervision, and are more likely to disengage once back at school due to the feelings of disconnection that can arise from classroom removal (Fenning et al., 2012; Skiba et al., 2014). Each of these can increase the risk of school dropout and justice system involvement.

It has been thought that zero tolerance policies--which heavily rely on exclusionary practices--played a large role in the creation of the STPP (Diem & Welton, 2020; Ruiz, 2017). Zero tolerance policies are “regulations that mandate specific consequences in response to outlined student misbehaviors, typically without any consideration for the unique circumstances surrounding a given incident” (Ruiz, 2017, p. 808). These policies increased in the 1990s in response to the Gun-Free School Act of 1994 as a way to combat concerns about school violence that seemed more extreme than in decades prior (Diem & Welton, 2020; Ruiz, 2017). These policies turned out to be an overreaction in comparison to the actual percentage of violent crimes committed in schools and an unintended consequence was the STPP (Ruiz, 2017).

Empirical research has repeatedly documented the ineffectiveness of zero tolerance policies (Cagle, 2017; Keleher, 2000; Ruiz, 2017). Over the past 40 years, suspension rates have doubled with one-third of all students experiencing some form of exclusionary discipline (Fisher et al., 2020; Hemez et al., 2020; NCES, 2020; Owens & McLanahan, 2020). In some of the more recent studies, it was documented that in the 2011-2012 school year alone, 3.5 million students received out of school suspension or expulsion (Owens & McLanahan, 2020). This is particularly concerning because of the negative outcomes correlated with school exclusion. In a study by Suh and Suh (2007), researchers found that a single suspension increased the likelihood of dropping out by almost 77% and that number of suspensions is more predictive of dropout than grade point average or socioeconomic status. With each additional suspension or expulsion,

there is a greater likelihood of a student dropping out, which is associated with a plethora of negative outcomes (Hemez et al., 2020). For instance, students who drop out have decreased employability and lifelong earning power which can lead to low income and further financial barriers that correlate with a greater need for social services. Additionally, they experience greater decreases in health, lower access to medical care, and an increased likelihood of engaging in criminal behavior (Ruiz, 2017). Each of the latter impact one's quality of life and create financial strain on society at large. That said, exclusionary discipline is in fact a systemic concern for society more broadly.

Despite the continued widespread use of exclusionary methods of discipline, there has been an increasing number of school districts transitioning to the use of newer research to inform improved discipline policies and practices. Harper (2020) reported on school districts in Texas and Pennsylvania that have transitioned to limiting the use of exclusionary practices for minor infractions and instead requiring disciplinary alternatives to be used prior to suspension. The author also showcased that, as of 2017, 27 states legally require the tracking of racial disciplinary discrepancies as a way to monitor and bring awareness to the racial discipline gap. However, they cautioned that, despite signs of progress, policy reform alone may not be enough to address the racial discipline gap. Many factors influence school discipline (e.g., staff culture, educational equity, perceived staff support, resources, staff and community demographics) and policy change may not be enough to address the nuanced interplay of these. Rather, person-centered approaches that focus on student and staff skill building, self-awareness, and social justice may also be needed.

### **Disproportionality and the School-to-Prison Pipeline**

School exclusion has been related to increased risk of incarceration later in life. Using the longitudinal data from the National Longitudinal Survey of Youth 1997, Hemez et al. (2020) explored the relationship between middle and high school student suspension(s) and incarceration in young adulthood (ages 18-26). The researchers analyzed suspension data for 8,984 individuals over a 14-year period (1997-2011). Using yearly surveys, they tracked number of suspensions, antisocial tendencies of peers, and contacts with the justice system. When doing so, they controlled for age, gender, marital status, level of education, mother's education level, household income, and criminal records. Results showed that among the participants who had never experienced suspension, less than one percent were incarcerated per year. Conversely, those who experienced at least one suspension between Grades 7 and 12 had increased odds of incarceration by 878%. That rate increased another 26% with additional suspensions, thus highlighting the potentially compounding effect of multiple suspensions. In similar studies with incarcerated adults and adolescents, 80% of adult and juvenile offenders had either dropped out of school or reported failing in school (Ruiz, 2017; Sander et al., 2011). Sander (2010) reported that almost a third of the individuals in juvenile residential facilities have experienced suspension or expulsion. The results of these studies beg several important questions. First, is suspension serving as a catalyst for the cascade of events that lead students to the justice system? Second, does justice system involvement occur because students are demonstrating higher levels of antisocial behaviors (resulting in suspensions) or does unnecessary suspension and/or expulsion help create conditions that creates or exacerbate undesirable behaviors? Third, and most importantly, are the higher rates of suspension and expulsion for Black youth reflective of differential rates of inappropriate behavior, or are there systemic variables (and if so, which

ones) contributing to these outcomes? These have proven to be difficult questions to answer, which has made disrupting the STPP challenging.

Further complicating these questions is how minoritized students (e.g., Black, Native American, Latinx, students with disabilities) are disproportionately impacted by the STPP (Fenning & Rose, 2007; Skiba, 2014). Compared to White peers, Black students are two to three times more likely to be suspended or expelled, Native American students two times more likely, and Hispanic/Latinx students 1.3 times more likely (Owens & McLanahan, 2020; U.S. Department of Education Office for Civil Rights [OCR], 2014). For over three decades, Black students have consistently been the most overrepresented population in school discipline and youth incarceration (Carter et al., 2017; DeMatthews, 2016a; Diem & Welton, 2020; Fenning et al., 2008; Gregory et al., 2010). This is particularly concerning because there has been no evidence to suggest that Black students engage in problematic behavior at higher rates than other student populations (Diem & Welton, 2020; Sander et al., 2011).

Disproportionality in exclusionary discipline for Black students has been shown to emerge as early as preschool, at which time these students are estimated to comprise only 18% of the student population but 48% of those suspended (Diem & Welton, 2020; Fisher et al., 2020; Graves & Howes, 2011; Owens & McLanahan, 2020; U.S. Department of Education Office for Civil Rights, 2014). More broadly, among students in K-12 settings during the 2013-2014 school year, Black students comprised 15.5% of students enrolled in schools nationwide yet represented 38.7% of those suspended (Fisher et al., 2020). Conversely, in that same year, White students comprised 50.3% of enrollments and represented only 32.5% of students suspended. In the 2015-2016 school year, similar trends were observed. From a sample of 89,323 students (46.5% White, 36.0% Black, 9.3% Hispanic, 8.1% other) from 131 public schools in the Midwest, 17.4%

of Black students were suspended at least once, a rate that was 3.8 times higher than White peers, even after controlling for gender, SES, grade, student-teacher ratio, and school location (urban city, large suburb, or rural area; Fisher et al., 2020). That said, not only are exclusionary practices correlated with extremely negative life outcomes, but they continue to disproportionately impact Black students.

Barnes and Motz (2018) conducted a longitudinal study that examined the racial discrepancy in which they used a multistage cluster sampling process to survey 20,000 students from approximately 100 schools. Sampling began in the mid-1990s and continued for 14 years. Sampling occurred twice while students were in middle or high school, once while transitioning to young adulthood (18-26 years of age), and once more as young adults (26-32 years of age). Results indicated a notable correlation between suspension and arrest rates for Black individuals even after controlling for common predictors of problem behavior in adolescents (i.e., poor relationships with teachers, involvement with substance-using peers, frequency of self-reported delinquency, self-reported drug use, level of self-control, level of depressive symptoms, level of attachment to parents, level of maternal involvement and education, verbal IQ, grade point average, neighborhood level of concentrated disadvantage, age, and sex). Not only were Black students almost twice as likely to be suspended or expelled, but those who were suspended were 60% more likely to be arrested compared to White peers--even after graduating from high school (Barnes & Motz, 2018). These findings help support the hypothesis that not only do exclusionary discipline practices contribute to justice system involvement, but that Black students are more susceptible to the STPP as a result. However, Barnes and Motz (2018) did note that an important limitation to their study was that even if exclusionary discipline policy changes were made, it is likely that a discrepancy in arrests for Black individuals would remain, which is a broader

societal issue beyond the scope of this study. Nevertheless, eliminating school policies and practices that are harmful and unevenly applied to minoritized students may go a long way in reducing the disparity in incarceration rates for Black youth.

### **Contributing Factors**

Because of the negative outcomes associated with the STPP, researchers have sought to find explanations for its existence. Sander et al. (2011) explored characteristics specific to individuals in hopes of informing student level interventions. Other student and family level factors that have been explored include the relationship between suspension of Black students and academic engagement, emotional engagement, negative or delinquent behavior history, behavior severity, mental health symptomology, student perceptions of school climate and culture, student perceptions of interactions with teachers, sense of belonging at school, SES, parental education, parental attachment, neighborhood instability, and peer delinquency (Diem & Welton, 2020; Fisher et al., 2020; Gregory et al., 2010; Jarvis & Okonofua, 2020; Novak, 2019; Sander, 2010; Sander et al., 2011). However, few studies have produced strong evidence that student level (aside from race) or family level factors are good predictors of the likelihood of justice system involvement (Diem & Welton, 2020; Gregory et al., 2010; Jarvis & Okonofua, 2020; Novak, 2019).

Some of the early efforts to understand the connection between race and exclusionary discipline practices were conducted by Gregory et al. (2010) by way of synthesizing existing research. The authors examined suspension and juvenile justice involvement trends and potential contributing factors among minoritized students. They reported that living in low-income neighborhoods where there was higher exposure to violence, substance use, and delinquent peers was predictive of suspension and justice system involvement for Black students. They also found

a relationship between GPA and race when predicting suspensions. However, the authors cautioned that unequal access to education and educational supports may better explain these findings, especially the relationship with GPA. Gregory et al. (2010) proposed that racial disparities likely begin at the classroom level, not the student level.

In a study by Novak (2019), data from the Longitudinal Studies on Child Abuse and Neglect (LONGSCAN) Child Protective Services dataset was used to explore whether there were relevant individual, family, or neighborhood variables that predicted school suspension. Data for 837 participants were gathered starting when children between four and six years of age and then every two to four years until participants turned 18. Factors controlled for included sex, race, SES, aggression, externalizing behavior, school commitment, neighborhood violence, learning problems, deviant peer association, and maltreatment history. Results indicated that receiving at least one suspension by age 12 doubled the likelihood of justice system involvement at later ages. Being male and associating with delinquent peers were also predictors of justice system involvement. Otherwise, all other factors were *not* significantly predictive of justice system involvement, suggesting the need to explore more systemic factors to help explain the racial discrepancy in the STPP.

Taken together, it is clear that there are no strong or consistent individual or family variables that predict the higher rate of school suspension among Black and other minoritized students. Although there are some predictive trends, such as race, being male, having a low GPA, and associating with delinquent peers, these same variables may also be explained by systemic considerations such as unequal access to quality early childhood education, living in unstable neighborhoods, staff skills and/or training, or staff perceptions of student characteristics. Therefore, a common theme among scholars who have researched student level characteristics is

the recommendation that future research focus on understanding systemic factors (i.e., classroom, administrative, and district level) that are associated with the racial discipline gap (Diem & Welton, 2020; Fenning & Rose, 2007; Gregory et al., 2010; Novak, 2019; Sander et al., 2011).

From an intervention perspective, focusing on student factors makes sense because implementing student-based interventions is likely easier than larger cultural paradigm shifts. This may be especially true regarding racial discipline discrepancies given the deep-seated history of systemic racial oppression in society at large as well as the slow pace of change in the education system. According to Bronfenbrenner (1979), all systems (individual, micro, meso, exo, and macro) play an important role in the development of individuals. As noted, there are few consistent correlations with suspension at the individual and family levels. Therefore, an evaluation of the interplay between micro and macro levels--classroom, school, and district--should be the next direction for future research.

### **School Policies**

A functional school system is dependent on the various policies used to guide day-to-day operations. Of these, discipline policies are of specific interest to this study. Such policies are intended to communicate behavioral expectations to students, school staff, and parents. The importance of having well-articulated discipline policies gained traction in the late 1970s when U.S. educational agencies began pushing for them (Fenning et al., 2012). The motivation behind this was preventive in nature such that by having a shared understanding of acceptable and unacceptable behavior, students would avoid punishable behaviors (in theory). In the 1980s, this approach morphed into zero tolerance policies as the Reagan administration declared its war on drugs (Law Library, n.d.). Zero tolerance policies were used to instate harsh, exclusionary



discipline to deter students from undesirable behavior by communicating that misbehavior of any level of severity would be punished (Novak, 2019). The national expectation became that if schools were to continue receiving federal aid, they would adopt zero tolerance policies, thus they quickly became the norm (Law Library, n.d.).

Unfortunately, zero tolerance policies did not achieve their intended goal largely due to issues with inflexibility and ambiguity (Diem & Welton, 2020; Fenning et al., 2012; Fenning & Rose, 2007). Adherence quickly led to disciplinarians suspending students for trivial or nonviolent offenses because it was outlined in their policies or discipline matrices to do so (Fenning & Rose, 2007; Law Library, n.d.). Fenning and Rose (2007) conducted a content analysis of secondary school disciplinary codes and discovered that suspension is the most common response to disciplinary infractions in high schools. They also reported that suspension and expulsion were the most frequently listed options for discipline consequences regardless of the severity of a behavior (mild, moderate, or severe). In a similar study, Fenning et al. (2008) analyzed data from the 2002 National Center for Education Statistics Directory where surveys were used to analyze 64 codes of conduct. Again, suspension and expulsion were the most frequently used consequence for all disciplinary infractions, regardless of severity. In a more comprehensive study, Fenning et al. (2012) conducted a content analysis of 120 school discipline policies across six states. The study analysis suggested that suspension and expulsion were assigned routinely for mild behavior infractions. Currently, about 95% of suspensions are for “nonviolent, minor disruptions” suggesting that zero tolerance policies continue to be used as a one size fits all response to discipline (Diem & Welton, 2020; Llorente, 2014; Ruiz, 2017). Taken together, it appears students continue to being subjected to exclusionary discipline practices for minor behaviors that were not initially part of the overarching goal of zero tolerance

initiatives (to decrease school violence). Despite this, zero tolerance policies are still commonplace with suspension continuing to be one of the most commonly used disciplinary consequences (Diem & Welton, 2020; Fisher et al., 2020; Gullo & Beachum, 2020b; Ruiz, 2017). An additional byproduct of zero tolerance policies that has also contributed to the STPP is the increase in school resource officers (SROs) and more inclusion of them in disciplinary processes. The increased law enforcement presence in schools has been associated with greater numbers of school-based arrests and juvenile detention referrals (Weisburst, 2019), which has also disproportionately impacted minoritized students who already receive more ODRs (Diem & Welton, 2020).

Another important factor to consider when explaining the high frequency of suspensions is ambiguity in decision making. When students end up in a principal's office for behavior concerns, it is usually via the route of a referral from a teacher. ODRs are usually made when staff observe students violating school rules/codes of conduct or when students have repeatedly broken classroom rules (Pas et al., 2011). When confronted with an ODR, principals must go through the process of determining whether a behavior concern warrants a consequence, and if so, what kind. Despite this process sounding straight forward, "if warranted" can be quite subjective depending on what kind of guidance a principal has been given by district leaders to navigate disciplinary decisions. For situations in which there is ambiguity or interpretive variability of the situation, examining the interplay between implicit biases of principals, ambiguity in discipline policies, and the effectiveness of cultural training of principals might help explain how disciplinary decisions result in racial discrepancies among Black and White students. Exactly how implicit biases, policy variability, and professional training might impact how principals navigate disciplinary ambiguity are outlined below.

## Bias

Broadly speaking, biases represent learned preferences. Much of human decision-making is at least partially guided by biases. There are many kinds (Ruhl, 2021), but the two most relevant to this study include explicit and implicit biases. Explicit biases are feelings, attitudes, and behaviors that are processed at a *conscious* level and recognized by an individual (National Center for Cultural Competence [NCCC], n.d.). The NCCC (n.d.) reported explicit biases are part of a reflective system in which thoughts and behaviors require intention and cognitive resources, which creates a cognitive load that can make such mental processes less efficient. Conversely, implicit biases are subconscious, pervasive, *automatic* mental processes that influence how people think about and respond to their environment (Carter et al., 2017; Ruhl, 2021). They develop from mental shortcuts called heuristics that are used to improve the efficiency of decision making, especially in complex, high stress, or repetitive situations (Fiske & Taylor, 2008). When considering everyday life, mental shortcuts are imperative for increasing cognitive efficiency; however, they are prone to errors and can develop into harmful thought patterns (Fiske & Taylor, 2008; Ruhl, 2021). The NCCC (n.d.) reported that unconscious biases are part of a reflexive system (in contrast to the reflective system) where long-standing information is stored and used to make decisions quickly so as not to create excess cognitive load. Because of the efficiency of the reflexive system, it is typically the primary processing route when an individual is stressed or experiencing high cognitive load.

Understanding implicit biases as they pertain to this study is important because they sometimes operate in opposition with someone's consciously stated beliefs and values (Carter et al., 2017; Gullo, 2020). A particularly relevant example of this pertains to "otherness", which is a type of heuristic that includes categories such as race, ethnicity, education level, or age--any way

in which a group or individual can be different than oneself (NCCCC, n.d.). From an evolutionary perspective, the ability to quickly discern out group members when humans lived in tribes served as a safety mechanism because “others” were more likely to be a threat (Fedor, 2014). This mechanism is much less needed in current society from the perspective of safety, yet it persists as it helps the brain categorize people in a way that makes it easier to decide with whom and how to interact based on previous experiences and learned preferences. A good example of such implicit biases may appear in help seeking behaviors. For instance, those who have experienced positive relationships with police and view them as trusted officials would be more likely to approach an officer to ask for directions as opposed to those who may have had negative experiences with them. In both instances, decision-making aligns with an individual’s preconceived beliefs or biases about an out-group member. This type of categorization system develops from generalized attributes and can lead stereotypes (Ruhl, 2021).

It is important to state that stereotypes are not inherently bad; rather, those rooted in systems of oppression are, especially in the education system. When considering the history of Black individuals in the United States, the potential harm of stereotypes towards Black males is especially relevant. Brown (2018) attributed harmful stereotypes about Black individuals to what he calls one of the worst “inventions of the twentieth century--the fabrication of the oppositional and dangerous Black male” (p. 54). Negative generalizations about these individuals were used to oppress and incriminate Black men and women for decades in ways that still permeate society. Brown (2018) credited the continuation of this legacy to “cultural memory”, the concept of how stereotypes and categorizations live on from one generation to the next even with explicit knowledge of their detriment. That these continue to exist is important to note because people can and do act in opposition to their stated beliefs and values (Carter et al., 2017; Gullo, 2020).

That is, an individual may believe in fair discipline practices, but when faced with a quick decision may make assumptions about who was at fault or choose a harsher penalty based on their own implicit biases (Smolkowski et al., 2016). For individuals working in high stress environments or those under high cognitive load, this is even more so the case due to an increased likelihood that mental shortcuts will be activated to increase efficiency.

Another aspect of otherness relates to the demographic makeup of principals and teachers in the United States. Approximately 80% of these individuals are White (Minkos et al., 2017; NCES, 2019, 2020a, 2020b), which may exacerbate the likelihood of harmful racial implicit biases. Thinking about the interplay between working in a high stress/high demand environment, deeply engrained societal Black stereotypes (Gullo et al., 2018), the race of those in power, and the function of the otherness heuristic, there is high potential for implicit biases to influence disciplinary decisions.

### **Existing Research on Principal Implicit Bias and Discipline**

Because principals are often the disciplinary gatekeepers, it is imperative to consider how their implicit biases influence their work with students. Very few studies have looked at principal implicit racial bias and its possible role in the racial discipline gap though (Jarvis & Okonofua, 2020). One of the first studies that included principal participants was completed by Gullo in 2017. The author researched whether school administrator implicit racial bias was predictive of discipline severity during subjective, objective, and overall discipline severity after controlling for student race, SES, severity of behaviors that led to a referral, student grade, administrator years of experience at their current site, and administrator race. Participants included 41 administrators from K-12 schools. It was found that implicit bias accounted for 89% of the variance of discipline severity for students of color in subjective discipline cases. No significant

effects were found for objective discipline scenarios though. The author cautioned that the low response rate (33%) could have potentially inflated such findings and thus encouraged further research.

In a very similar study, Gullo and Beachum (2020a) used a race IAT to explore the possible impact of implicit bias on discipline outcomes with 43 principals. Of the participants, 67% showed a slight preference for white students. Results also showed implicit bias accounted for discrepancies in consequence severity between White and minoritized students during subjective but not objective discipline scenarios. For instance, when discipline referrals contained behavioral ambiguity (e.g., misconduct, disrespect, insubordination), principal implicit bias explained the difference in discipline severity between Black and White students. The severity of discipline between Black and White students during objective disciplinary scenarios (e.g., fighting, bringing a weapon to school, possession of drugs) however, were not better explained by implicit bias.

An additional study by Jarvis and Okonofua (2020) also explored how disciplinary decisions might be impacted by student race. The authors recruited 91 assistant principals to answer questions about fictitious discipline scenarios with two different students. Each student had either a stereotypical Black or White name and was referred to the office on two separate occasions. Participants responded to a set of questions after the first referral (i.e., possible characteristics of each student, severity of the behavior, how irritated the participant felt, how severely they would discipline the student). The principals then read a second referral for an incident that occurred three days later for the same two students and responded to similar questions. Results indicated that the Black student's behaviors were rated as more severe and they were assigned more severe consequences for both the first and second referral. In their

concluding remarks, the authors hypothesized a possible reason for their findings. Historically, existing research has shown that teachers discipline Black students more harshly than White students only after repeat misbehavior. In contrast, the assistant principals assigned harsher consequences upon the first referral. When examined together, one might conclude that teacher racial biases likely play a role in the *start* of the disciplinary process; however, the role of principal implicit bias in the discipline hierarchy has remained less clear. The authors questioned if principals tended to be harsher because they trusted that teachers only referred students they no longer felt equipped to manage (thus, potentially justifying a more severe consequence upon an initial referral) or if principal racial implicit biases may be at play. Given the author's findings, the latter appeared to be true in their study.

Aside from the formerly discussed studies, very little quantitative research currently exists on principal implicit biases and the disciplinary process. To date, the majority of empirical studies on implicit racial biases in schools have been conducted with teachers (Carter et al., 2017; DeMatthews et al., 2017; Graves & Howes, 2011; Gullo & Beachum, 2020a; Sander et al., 2010; Skiba & Leone, 2002). Fortunately, studies exploring the correlation between teacher biases and student academic success may help inform the design of future studies with principals. In one study on the influence of teacher implicit bias, explicit bias, and their expectations of students on math and reading test scores, Van den Bergh et al. (2010) assessed potential differences in achievement scores between “non-preferred student groups” (i.e., ethnic minorities) and White students. To measure bias, teachers completed an implicit association task (IAT). The study resulted in two key findings. There was no significant correlation between teacher *self-reported* bias and their expectations of student achievement; however, teacher ratings on the IAT helped explain the achievement gap for ethnically minoritized students. The

key takeaway was that implicit bias may be present even when teachers self-report low levels of it, meaning that self-report measures alone seem to not be enough to identify negative effects of implicit racial bias. The results also indicated that implicit bias and its role in staff expectations of students can influence student experiences at school in ways that do not pertain to the disciplinary process but that can contribute to the feelings of disconnection that may lead to negative behaviors (Gullo, 2020).

In a study more specific to race and discipline, Girvan et al. (2016) explored the significance of implicit bias in subjective discipline scenarios by looking at over one million ODRs across 1,824 schools. They found subjective ODRs explained 1.5-3 times the variance of racial disproportionality between Black and White students in comparison to objective ODRs. These results suggest that disproportionality in exclusionary discipline might be attributable to implicit bias to some degree when there is ambiguity in disciplinary decisions (Girvan et al., 2016). Curiously, in a study by DeMatthews et al. (2017) that examined the nature of ODRs, the authors reported that Black students engaged in less severe behaviors than White students yet received more ODRs than any other racial group. They also consistently received harsher punishments than White peers who committed the same infraction type (Barnes & Motz, 2018; Carter et al., 2017; Diem & Welton, 2020; Fisher et al., 2020; Gregory et al., 2017; Skiba, 2014). Through the lens of bias, disciplinary issues may be more a matter of staff perception of students than a question of behavioral severity. Owens and McLanahan (2020) investigated this hypothesis as well and reported that discrepancies in teacher perceptions of behaviors based on student race accounted for 46% of the discipline gap between Black and White elementary students. Conversely, differences in actual behavior severity accounted for only 9 percent of the gap. In response to these results, the authors questioned whether Black students truly engage in



more problematic behavior or if staff racial biases create the illusion that they do. Additionally, teacher biases may also help explain the higher rate of behavioral referrals for Black students.

Despite studies highlighting how teacher implicit bias clearly plays a role in the increased likelihood of Black students being referred to the office, the racial discipline gap persists beyond them as can be seen in the ongoing racial discipline gap (Gullo & Beachum, 2020a; Gullo et al., 2018; Jarvis & Okonofua, 2020). Because suspensions fall under the jurisdiction of the principal, a reasonable next step is using the studies that have been conducted with teachers to help inform future research on the role of principals' implicit bias in the disciplinary process (Gullo & Beachum, 2020a; Jarvis & Okonofua, 2020; Williams et al., 2020).

### **Bias and Subjectivity**

As previously alluded to, the impact of implicit racial bias is especially important in the context of subjective versus objective disciplinary decisions. Objective decisions are straight forward due to concrete behavioral definitions that guide their interpretation. For example, there is no ambiguity or interpretable variability in bringing a weapon to school, physical fighting, or possessing drugs. Conversely, subjective disciplinary scenarios include interpretive ambiguity. Examples of these have included behaviors such as disruption, defiance, misconduct, disrespect, or insubordination--all of which can be culturally laden terms and interpreted differently across situations, individual staff, and districts (Pinto, 2013). In such situations, when information is ambiguous, mental shortcuts (i.e., biases) will help fill in any gaps (Greenwald & Krieger, 2006). When someone is under a high degree of cognitive load--as principals often are given the fast paced and multifaceted nature of their roles--efficiency becomes a priority. Unfortunately, when principals are short on time, information, and cognitive resources, their decision making is more susceptible to negative implicit biases (Gullo et al., 2018). Regarding discipline decisions with

Black students, this would be a time when stereotypes are much more likely to influence decision making if implicit racial biases go unchecked. Sadly, even for those individuals who report being committed to equality, implicit bias can still counteract self-reported beliefs (Gullo et al., 2018). Simply being in the presence of someone of a race other than one's own is enough to elicit stereotypical thoughts and trigger otherness heuristics, regardless of whether they align with one's own stated beliefs (Brown, 2018; Gullo et al., 2018). This combination of factors can work against even the most student-oriented principals with the best of intentions, which begs the question of what is to be done about the issue of implicit racial biases in discipline then. In order to conduct further research to help answer that question, a basic understanding of principal training and preparation was required.

### **Principal Preparation Programs and Professional Standards**

In this study, educational leadership is used as a broad term to refer to any school personnel involved in disciplinary decision-making processes including, but not limited to principals, vice principals, assistant principals, and deans. That said, principal education and preparation have come a long way since early practice when principals were trained more like corporate business managers (Murphy, 2005). As a result, principals were not student- or community-centered, they were business oriented, which resulted in an educational research to practice gap (Farley et al., 2019; Smylie & Murphy, 2018; Williams et al., 2020). In an effort to remediate this and reform the role of principals, educational leadership standards were created. The first version came from the Interstate School Leaders Licensure Consortium (ISLLC) in 1996 and was written for *practicing* principals (Farley et al., 2019). However, there was initially no corresponding set of standards for principal preparation programs to help align training curriculums with practice standards. Thus, a discrepancy began to form between training

received and the applied skills needed to serve rapidly changing school systems (Farley et al., 2019). To combat this, the Educational Leadership Constituent Council (ELCC) developed standards for preparation programs in the early 2000s that closely aligned with ISLLC standards to help reduce training-to-practice misalignments and with the intent to move the field towards a more scientist-practitioner model (Farley et al., 2019; M. Young, 2018).

With continued efforts to foster student- and community-centered principals, the ISLLC standards were replaced in 2015 by the Professional Standards for Educational Leaders (PSEL; National Policy Board for Educational Administration [NPBEA], 2015) and the ELCC standards were replaced with the National Educational Leadership Preparation (NELP) program standards (NPBEA, 2018). Both updates were more equity and student-centered than previous versions. The PSEL specifically called for culturally responsive practices to support the inclusion and success of all students (Minkos et al., 2017). However, many researchers, educators, and practitioners have reported concerns that neither of these updates provide comprehensive practice guidelines on how to be an equitable educational leader (Farley et al., 2019; Smylie & Murphy, 2018; M. Young & Perrone, 2016). Although there was a notable increase in equity focused language in both editions, the terms such as equity, equality, social justice, at risk, race, and cultural competence were not explicitly defined or contextualized (Farley et al., 2019). Furthermore, the revised editions did not overlap in a way that aligned pedagogy and practice, continuing to leave practicing leaders without some of the knowledge or skills they needed to be equitable and/or effective after leaving their principal preparation programs. More importantly, few states have adopted the PSEL standards, acting as a further barrier to culturally responsive leadership. Since their release, only 12 states have endorsed them (Smylie & Murphy, 2018).

Most states still use the ISLLC standards to “recruit, develop, and evaluate” principals and school systems (Swearingen, 2014, p. 21).

There is ample support for having educational leadership standards for practicing professionals and training universities, but so far, there is limited research examining their impact and how they may be influencing more culturally responsive practices (Farley et al., 2019; McCarthy et al., 2016; Smylie & Murphy, 2018; M. Young & Perrone, 2016). With the slow adoption of the PSEL standards across the country, it has been difficult to evaluate their utility and incorporation into training programs (DeMatthews, 2015). Many principals have expressed a need for training that equips them with actionable behaviors to lead in equity focused ways (B. L. Young et al., 2010). Currently, there are mid- to late-career educational leaders who may have had minimal formal education on the interpretation and implementation of the new standards. With the expectation for principals to use models of equity-oriented leadership, it is important to explore what types of training--both in preparation programs and beyond--create leaders who are sufficiently prepared to lead schools in equitable ways, which would include awareness of implicit racial biases as encouraged by the PSEL standards (Swearingen, 2014).

### **Principal Preparation Programs**

Currently, researchers, educational leadership faculty, and practicing principals have expressed concern that graduate programs may not be training students to lead schools in culturally responsive ways (Minkos et al., 2017). Khalifa et al. (2016) described culturally responsive school leadership as focusing on anti-oppressive, anti-racist leadership that consists of “practices and actions, mannerisms, policies, and discourses that influence school climate, school structure, teacher efficacy, or student outcomes” (p. 1274). Regarding school leadership, research

on how cultural responsiveness is developed is scarce. Little is known about how preparation programs are offering opportunities for their students to develop knowledge, awareness, and skills related to culturally responsive practices. Of the studies that do exist, many lack a robustness that would allow for generalization.

One example came from Miller and Martin (2015) who conducted a qualitative study of four principal's perceptions of how their training programs and professional developments prepared them for addressing social justice issues in schools. The participants were from an urban district, had at least five years of experience, consisted of one male and three females (two were White and two were Black), and were placed at all levels of school (i.e., elementary, secondary). Participants reported neither their training programs nor subsequent professional development prepared them with actionable ways to address social justice issues. All participants reported being equity-oriented leaders but were often stuck in a "deficit thinking paradox", pictured in Figure 1. This paradox referred to personal blind spots and/or biases that emerged when participants offered answers that contradicted with what it means to be equity-oriented even when describing themselves as such. For example, participants talked about closing the achievement gap and increasing standardized test scores as a means of addressing inequity but did not seem to convey that diversity is a whole child process. Also, participants did not describe how systemic changes, such as shifting the cultural mindset of all school staff, could alleviate the achievement gap as well as larger social injustices. Other researchers have noted this trend as well and expressed that oppression is most likely to occur when leaders hold deficit-oriented opinions (Khalifa et al., 2016; Minkos et al., 2017).

**Figure 1**

*The Paradox for the Principals Understanding Social Justice*



*Note.* Reprinted from “Principal Preparedness for Leading in Demographically Changing Schools: Where is the Social Justice Training?” by C. M. Miller & B. N. Martin, 2015, *Educational Management, Administration, & Leadership*, 43(1), p. 147. Reprinted with permission (see Appendix A).

In an effort to explore the content of principal training programs as related to culturally responsive instruction, Cox (2017) evaluated the content of four programs in South Carolina. The author reviewed course syllabi, program specific documents, internship requirements, and conducted interviews with program coordinators. They found that only one program aligned its curricula and experiential opportunities in a way that would be consistent with culturally responsive training. Cultural responsiveness content was missing from almost all courses from the remaining three programs. Instead, training programs primarily focused on ELCC Standard 3, which focuses on school management and organization (NPBEA, 2011). Results also indicated there were few, if any, opportunities to engage in applied experiences with culturally diverse populations. Few classroom assignments asked students to engage with mock analyses specific to potential culturally sensitive data, practice solving equity or cultural diversity

problems, practice promoting positive school culture, or receive instruction on how to choose and/or conduct professional developments specific to diversity issues. In summarizing their findings, Cox (2017) concluded that “students were not provided the content, instructional strategies, assessment, and field experience to acquire the skills to respond to culture and diversity issues as a school leader” (p. 299).

Although these studies are small and qualitative in nature, they are important for two reasons. First, they echo growing sentiments of scholars who have expressed concern regarding lack of training and opportunities to grow as culturally responsive school leaders (Marchioli et al., 2020; Moughania, 2018). Second, they also showcase the need for further research on principal training as it pertains to diversity. Related to the current study, robust training in this area may have the potential to offset negative implicit racial biases that might be contributing to the racial discipline gap. There is evidence that the detrimental impact of negative implicit biases can be altered with training that facilitates increased awareness of implicit biases (Lai et al., 2016). Changing implicit biases also takes repetitive exposure to training material though (Lai et al., 2016), but to what extent is still unknown in the context of principals and disciplinary decision making.

### **Continuing Education and Practicing Principals**

Looking at the professional world beyond preparation programs, continuing education is often the source of training for fostering cultural responsiveness once school leaders have entered the field. In general, diversity trainings aim to address negative attitudes toward minority groups by providing information about the history, experiences, and cultures of different groups of people as a means of fostering understanding and empathy through the use of perspective taking (Herbstrith & Busse, 2020; Jackson et al., 2014). Defined by Lindsey et al. (2014),

diversity trainings constitute “any program designed to facilitate positive intergroup interaction, reduce prejudice and discrimination, and generally teach dissimilar others how to work together effectively” as a means of dismantling systems of oppression (p. 606). Currently, there is lack of consensus about whether or not diversity trainings are effective and, if they are, what exactly makes them so (Bezrukova et al., 2012, 2016; Hussain et al., 2016; Jackson et al., 2014; Lindsey et al., 2014; McIntosh et al., 2021; Noguera, 2009). However, emerging research has suggested that the amount and type of trainings is helpful in building the skills needed for cultural responsiveness (Barakat et al., 2021; Hall & Theriot, 2016; Lindsey et al., 2014). Additionally, using social-psychological theory, Herbstrith and Busse (2020) posited that implicit biases do not need to be eliminated or unlearned for behavior change to occur. Thus, rather than viewing diversity training as a tool for bias elimination, it could be viewed as a method for shifting behavioral norms by creating the “normative pressure” needed for behavior change (Herbstrith & Busse, 2020). Said another way, the authors offered the opinion that diversity trainings may assist in creating new expectations across the professional world. As new norms are established, individuals who have been made aware of these via diversity trainings are more likely to behave in accordance with such norms in order to be viewed as socially acceptable.

In support of the effectiveness of diversity training, in a narrative review of 178 diversity training articles, Bezrukova et al. (2012) found organizations and educational programs that implemented an integrated model of diversity training (where diversity conversations were embedded across many parts of a job such as supervision, coursework, research, writing, marketing) were viewed more positively than those offering only one-time trainings. Such a claim was made based on participants’ valuations (post- course/training evaluations) of the perceived effectiveness of integrated approaches versus standalone trainings. These findings



imply that greater exposure to diversity training across different contexts may produce a greater sense of preparedness and generalization in diversity work. Unfortunately, the authors found that the bulk of diversity training studies to date have focused on stand-alone trainings.

In a longitudinal study examining the impact of diversity trainings, Lindsey et al. (2014) used a 3x2 pre-post design to examine 118 college undergraduates' attitudes of two minoritized groups (either Black individuals or sexual minorities). Participants were randomly assigned to one of three diversity training conditions and then randomly assigned to one of the two minority population conditions. Training types included guided perspective taking, goal setting, or stereotype discrediting. After receiving a diversity training during freshmen orientation, two check ins were conducted over an eight-month period using either the Modern Racism Scale or the Attitudes Towards Lesbians and Gay Males Scale. Researchers reported that despite participant reports of self-perceived improvement in diversity skills, the results of the administered scales suggested no evidence of behavior change. This could be in part due to study limitations (i.e., sample type, limited heterogeneity, high attrition), but it could also speak to the need for more research beyond standalone training experiences. Despite variability in the results of diversity training studies, there is empirical data to suggest that they have value. However, there is still the question of whether the most effective training is the result of repeated, integrated opportunities for learning--which is not currently the norm for ongoing professional development in school settings.

### **Building Culturally Responsive Practitioners**

Basic learning theory and behavioral psychology have both been used to suggest that repetition is a key aspect of effective learning and generalization (Hartley, 1998). Assuming this applies to principal preparation and continuing education trainings, it could be hypothesized that

the more diversity training hours a principal engages in, the more knowledge they are equipped with and the higher the chances of them generalizing their knowledge across situations. As this hypothesis pertains to implicit racial bias, it could be surmised that with appropriate *ongoing* training, principals may gain greater awareness of their own implicit biases and be better equipped to monitor them when interacting with students from minoritized groups. Forscher et al. (2019) supported this concept, expressing that “repeated pairings” of newly acquired information with relevant social contexts is needed for meaningful, sustained change to happen (p. 542).

The value of greater amounts of training was further supported in the work of Hall and Theriot (2016) who studied cultural competence in 191 social work students from three universities. The researchers used a pre-posttest design to compare diversity training across different locations, course instruction, and assignments to see how each influenced student cultural competence. All students were required to be enrolled in at least one course with a diversity component to qualify for participation. Prior to taking these courses, participants completed the Multicultural Awareness-Knowledge-and Skills Survey (MAKSS) and described previous diversity training experiences they had had in other settings (e.g., prior education courses, employment, church) to gauge each student’s preexisting knowledge of diversity. At the end of the semester, students again completed the MAKSS. At post-test, all test scores markedly improved; however, the only consistent predictor of total point increase was the number of diversity trainings/settings *prior* to starting the semester. The more diversity training exposure participants had prior to the course, the higher their post-test scores were. For each additional site at which a participant had received prior training, their MAKSS post-test scores increased by five points.

Few studies have examined this concept with principals. In a mixed-methods study of 46 students attending a combined Master's degree and educational leadership certification program, Barakat et al. (2021) explored the perceptions of participants receiving a "cross-curricular" program. These individuals were the first cohort to receive a revamped curriculum at their program that was designed include components of social justice across *all* of their classes. To measure cultural competence, researchers administered the Cultural Competence for Educational Leaders (CCEL) instrument at the beginning and end of the program. The CCEL has three subtests that gauge cultural knowledge, cultural skills, and cultural beliefs and motivations as well as an overall cultural competence scale. Large to moderate effect sizes were found for all four scales at posttest (knowledge, 1.082; skills, .920; beliefs/motivation, .548; overall cultural competence, 1.094). The authors concluded that using a training approach that embedded social justice and cultural competence concepts across all parts of the program (i.e., during instruction and applied experiences), thereby increasing exposure and repetition, was associated with significant change on the subscales of the CCEL. Results suggested that repetition of concepts and opportunities for continued learning experiences may be a key part of developing culturally responsive leaders.

Because many training programs for educational leaders are relatively brief (approximately two years), it is imperative that training on cultural responsiveness continue after graduation. Khalifa et al. (2016) noted that preparation programs focused on developing culturally responsive leaders are necessary, but that is merely a foundation. Through their results, Barakat et al. (2021) helped demonstrate that integration of social justice elements throughout the curriculum may inform actionable steps for redesigning principal preparation programs. The question of how to foster cultural responsiveness for principals who are already practicing still

remains though. Those already in the field may not have had exposure or opportunity to engage in integrative or repetitious trainings that would foster cultural responsiveness. Just as principal preparation programs are in need of greater focus on social justice and culturally responsive practice, there is also a need to ensure that practicing principals are required to participate in ongoing training that enhances their skills as culturally responsive leaders *and* disciplinarians. An important first step is to explore whether there is a relationship between engagement in diversity trainings, principal racial implicit biases, and the racial discipline gap.

### **Summary**

Black students continue to be overrepresented in both exclusionary discipline and the juvenile justice system even after accounting for individual, family, and teacher level factors. With principals being the disciplinary gatekeepers, it is a reasonable next step to explore their role in the STPP. One way of doing so is exploring how their implicit biases may contribute to differential disciplinary treatment for students of different races. Additionally, it is also important to simultaneously investigate whether the amount of diversity training a principal has received modulates suspected negative implicit biases. In doing so, it may be possible to inform training that better supports principals and guides needed policy changes.

### CHAPTER III

#### METHODOLOGY

Existing research has shown there is a longstanding racial discipline gap between White and minoritized students, especially Black students. However, little research currently exists on how principals' implicit racial bias may contribute to this gap, especially as it pertains to subjective discipline. Whereas objective discipline refers to concrete definitions of problematic behavior, subjective discipline refers to behaviors that are subject to interpretive variability and are thus more prone to the influence of implicit biases (Girvan et al., 2016; Greenwald & Krieger, 2006; Gullo, 2017). Therefore, the purpose of this study was to explore how high school principal's implicit racial biases may influence the severity of consequences given to Black students during subjective discipline situations. This study also explored whether higher quantities of diversity training might offset or mitigate the impact of implicit racial biases. The design of the study was based on the hypothesis that higher numbers of diversity training hours would correlate with lower discipline severity for Black students, even if participants demonstrated higher pro-white preferences. This study used a non-experimental design via an online survey that included the collection of demographic information, disciplinary vignettes, and a race implicit association task (IAT).

#### **Instrumentation**

Three different types of measures were used for data collection, all of which were administered using Qualtrics. The order of instruments included demographic information, discipline vignettes, an estimation of the number of hours of diversity training, and a name-based

race IAT. This order was specifically chosen as a way of reducing social desirability, priming effects, stereotype threat, and survey attrition (McGlone & Pfiester, 2007). Demographic questions, vignettes, and the IAT are included in Appendices B, C, and D, respectively.

### **Demographic Information**

Respondents were asked to provide basic information about themselves from provided answer banks. These included their race (Asian, Black/African American, Hispanic/Latinx, Native American/Alaskan Native, Native Hawaiian/Pacific Islander, White/European American; Wisconsin Department of Public Instruction, n.d.), gender (female, male, non-binary/third gender, prefer not to say), current job title (principal, assistant principal, vice principal, dean), number of years worked at their current school (0-3, 4-6, 7-9, 10+), types of discipline used in the past year (restorative practices, detention, in school suspension, out of school suspension, expulsion, corporal punishment), and an estimate of the number of hours of diversity training completed in the last three years (0-3, 4-6, 7-9, 10-12, 13+). Diversity training was allowed to have occurred in higher education classes, professional development trainings, or as continuing education credits. A definition of what constitutes diversity training was provided for participants. Based on the literature review conducted for this study, the definition provided was, “any type of instruction that intended to foster cultural responsiveness, provide information on effectively interacting with diverse populations, provide information on implicit and/or explicit biases, encourage self-reflection of personal biases and/or personal cultural identity/intersectionality, reduce prejudice and discrimination, challenge preexisting knowledge about others’ experience and identities, provide information on how to encourage positive group interactions with dissimilar others, and/or provide information on dismantling systems of oppression”.

### **Vignettes**

To investigate the potential discrepancies in consequence severity given to Black versus White students, participants were asked to respond to three subjective disciplinary vignettes that were created via consultation with an expert panel of educational leadership professionals. The panel was used to ensure vignettes were subjective, differed in behavioral severity, and that they did not reflect any possible racial biases of the researcher. The panel also provided feedback on the names to be used in the vignettes. Three names were needed including one stereotypical White name (Brian), one stereotypical Black name (Darius), and one name that was common across races/ethnicities (Jacob; Sisense, n.d.; Social Security Administration, n.d.). Additionally, panel members provided feedback on the discipline consequences participants would be asked to choose from. They provided feedback and suggestions to ensure discipline consequence options were realistic and increased in severity/exclusiveness. During a brief pilot study, it was also determined via participant feedback that the discipline options provided seemed restrictive and did not allow for restorative approaches. For this reason, an additional discipline consequence option (refer for tier 2 supports and develop a behavior support plan) was added that aligned with participant feedback. Each response option was then quantified such that one was least severe and seven was most severe. Each of the vignettes and consequence options are listed in Table 1. The expert panel consisted of one female educational leadership professor from the Eastern United States who identifies as Black, one female educational leadership professor from the Western United States who identifies as White, and a bilingual public education director from the Western United States who identifies as a Latinx male.

**Table 1***Subjective Vignettes and Consequence Options*

Vignettes	Consequence Options
[Name] was talking during independent work time. When I asked him to stop because I should not be able to hear him all the way across the room, he told me to “just ignore it then, everyone else is”.	No consequence Parent phone call Detention
This is the second time this week [Name] has told me an assignment is stupid and not worth his time. I’ve asked to talk privately about what’s getting in the way and maybe to talk about alternative options for him and he just says the work is stupid in a rude tone.	Refer for tier 2 supports and develop a behavior support plan In school suspension
[Name] slammed his book down during independent reading and yelled, “this is effing stupid!”. I reminded him of how much time was left (just 5 minutes) and to be respectful of his classmates until then. He put headphones in and turned his music up loud enough that the whole class could hear. This is my fourth interaction like this with him this semester.	Out of school suspension Recommend for expulsion

*Note.* Vignettes are listed in order of increasing behavior severity. Consequence options are also listed in order of increasing severity.

Once the vignettes, names, and consequences were finalized, three separate versions of the survey were created to produce randomization across participants. The only difference between the surveys was which name was included in which vignette. For example, in each Qualtrics survey all demographic questions, vignette wording, and implicit association task were the same; however, Brian, Darius, and Jacob were rotated into different vignettes across the three different survey groups.

### **Implicit Association Test**

To measure racial implicit bias, a name-based race IAT was administered. IATs are reaction-based tests used to measure implicit biases. The original IAT method was developed by



Greenwald et al. (1998) in the 1990s and has since become popular in the social sciences due to Harvard's Project Implicit (n.d.). They have been used as alternatives to explicit tests of bias due to the latter often being more prone to social desirability effects (Stoet, n.d.).

To create a race IAT for this study, the original stimuli from the Greenwald et al. (1998) study were used (see Table 2). Initially, participants were going to be asked to complete the task using a platform called PsyToolkit. The program offers pre-programmed psychological experiment activities (Stoet, n.d.), meaning stimuli are already formatted for research administration, including all scoring procedures. However, when attempting to customize the IAT such that participants would not receive feedback on their biases (e.g., strong preference for White individuals, moderate preference for Black individuals), the researcher was unable to successfully alter the testing code to remove the feedback screen at the end of the administration. An Institutional Review Board (IRB) amendment was submitted and approved to make this change under the condition that resources about implicit biases were included at the end of the survey. An additional error with the platform was also encountered when attempting to calculate the D scores of completed test administrations; scores that were well out of range of the Greenwald et al. (2022) scoring formula were consistently being produce. For this reason, the IAT was instead created using an IAT generator called IATGEN that was compatible with Qualtrics (Carpenter et al., 2019), rather than PsyToolkit. Doing so allowed the researcher to ensure participants did not receive feedback about their level of implicit bias at the end of the task and allowed for adherence to the Greenwald et al. (2022) scoring procedure.

**Table 2***Implicit Association Test Stimuli*

White Names (target A)	Black Names (target B)	Pleasant Words	Unpleasant Words
Adam	Alonzo	caress	abuse
Alan	Alphonse	cheer	accident
Andrew	Darnell	diamond	assault
Brad	Deion	diploma	bomb
Brandon	Everol	family	cancer
Chip	Jamel	freedom	crash
Frank	Jerome	friend	death
Fred	Lamar	gentle	disaster
Greg	Lamont	gift	divorce
Hank	Lavon	happy	evil
Harry	Lerone	health	filth
Ian	Leroy	heaven	grief
Jack	Lionel	honest	hatred
Jed	Malik	honor	jail
Jonathan	Marcellus	laughter	kill
Josh	Percell	love	murder
Justin	Rasaan	loyal	poison
Matthew	Rashaun	lucky	pollute
Paul	Terrence	miracle	poverty
Peter	Terryl	paradise	rotten
Roger	Theo	peace	sickness
Ryan	Torrance	pleasure	stink
Stephen	Tyree	rainbow	tragedy
Todd	Tyrone	sunrise	ugly
Wilbur	Wardell	vacation	vomit

As mentioned above, the premise of an IAT is rooted in reaction time and assumes that individuals will more quickly associate compatible stimuli than they would incompatible stimuli.

For instance, on a race IAT using White and Black names, if an individual demonstrates a

stronger preference for White individuals, they will more quickly associate White names and positively valenced words and vice versa for Black names and positively valenced words.

Conversely, they will more slowly associate White names and negatively valenced words and vice versa for Black names and negatively valenced words. This is considered to be due to the mental effort required to “override [preexisting] mental associations”, thus taking individuals longer to complete associations (IATGEN, n.d., para. 6).

Of note, in a meta-analysis discussed by Greenwald et al. (2022), the test-retest reliability of IATs completed by an individual on one single occasion has an  $r = .50$  with internal consistencies of  $\alpha = .80$ . When interpreting these findings, the authors concluded that without multiple individual IATs to average and interpret, it is the aggregate  $\alpha = .80$  that best explains the systematic variance of IATs. To meaningfully interpret the reliability of one individual’s implicit bias, the authors recommended participants complete multiple administrations of an IAT. Repeat administrations of the IAT for participants was not feasible given the design of this study, thus the interpretation of the first research question is reflective of aggregate data.

During administration of the IAT, participants were asked to complete the standard seven block procedure specified in Greenwald et al. (2022; see Table 3), which took approximately 5 minutes. Illustrated examples of the IAT procedure used for this study are included in Appendix C. Participants were first acclimated to the task during blocks 1 and 2 during which they sorted target stimuli (Black or White names) and category stimuli (pleasant or unpleasant words) separately. During all blocks, participants were provided immediate corrective feedback if a name or word was not sorted into the accurate category. Black/White names appeared in the upper right or left corners of the screen in capital letters and pleasant/unpleasant words appeared in the upper right or left corners of the screen in lowercase letters regardless of the block number.

**Table 3***Standard Seven-Block Procedure*

Seven-Block Procedure	Recommended Number of Trials
Block 1: Classify the items for the two target categories	20
Block 2: Classify the items for the two attribute categories	20
Block 3: Classify items for all four categories, on attribute and one target category assigned to each of the two keys, using the assignment of categories to left and right keys as in Blocks 1 and 2	20
Block 4: Same as block 3	40
Block 5: Classify the two target categories, reversing the key assignments of Block 1 and having more trials than in Block 1	30
Block 6: Classify items for all four categories, using the reversed key assignments of the target categories as in Block 5	20
Block 7: Same as Block 6	40

*Note.* Standard seven block IAT procedure as outlined in Appendix A of Greenwald et al. (2022).

Participants then progressed to combined blocks during which they responded to *both* targets and categories (blocks 3, 4, 6, 7). An additional practice block was included between combined blocks 4 and 6 due to targets being presented in reversed position. Because the target locations switched corners in blocks 6 and 7, this additional practice block allowed participants to adjust to new target positions. By doing so, results could be considered reflective of likely biases rather than prior learning interference (Carpenter et al., 2019). Across all blocks, when stimuli appeared on the screen participants were asked to sort them as quickly as possible using either their left hand to press the “E” key or their right hand to press the “I” key. They were

provided immediate corrective feedback on all blocks if a word or name was not sorted into the accurate corresponding category. The key associated with each target/category was indicated on the directions page prior to every block administration. To ensure the task was completed in such a way that participants correctly used the “E” and “I” keys, they were required to use a computer or laptop with a QWERTY keyboard. To control for this, all three versions of the Qualtrics survey were programmed such that individuals using mobile devices (i.e., smartphones, tablets, iPads) were redirected to a page informing them that the survey could not be completed on a mobile device when attempting to access the survey link.

The left and right starting positions for targets and categories were automatically counterbalanced via IATGEN programming to create four possible starting configurations (White names on the right initially positive [compatible]; White names on the right initially negative [incompatible]; White names on the left initially positive [compatible]; White names on the left initially negative [incompatible]) that were randomly assigned across surveys (Carpenter et al., 2019). All scoring procedures were automated in IATGEN and aligned with the D measure scoring algorithms provided in Appendix B of the Greenwald et al. (2022) article, which produces a numeric estimate of implicit bias (see Table 4).

**Table 4***D Score Quantitative Categories and Qualitative Descriptors*

D Score Ranges	Qualitative Descriptor
> 0.65	Strong preference for White names
0.36 - 0.65	Moderate preference for White names
0.15 - 0.35	Slight preference for White names
0.15 - (-0.15)	No preference
(-0.16) - (-0.35)	Slight preference for Black names
(-0.36) - (-0.65)	Moderate preference for Black names
< (-0.65)	Strong preference for Black names

*Note.* Descriptors as seen in Haider et al. (2011).

### Procedure

Prior to recruiting participants, permission to conduct this study was obtained from the University of Northern Colorado's Institutional Review Board (see Appendix E). All instrumentation was administered via Qualtrics. The first page of the survey described the study, outlined inclusion criteria, provided a brief explanation of how data were to be used, and the researcher's contact information. Participants were informed that by choosing to proceed with the survey, they were providing their informed consent.

### Research Questions

- Q1 After accounting for hours of diversity training, is principal implicit bias a significant predictor of discipline severity on Black student vignettes?
- Q2 Do differences exist between the severity of consequences assigned to Black and White student vignettes?
- Q3 Do differences exist between the severity of consequences assigned to Black and White student vignettes accounting for hours of diversity training?

### **Data Analysis**

The following statistical analysis were used for each research question:

- Q1     After accounting for hours of diversity training, is principal implicit bias a significant predictor of discipline severity on Black student vignettes?

To best understand how each independent variable predicts discipline severity on Black student vignettes, a hierarchical multiple regression was used to examine whether principal implicit bias was a significant predictor of discipline severity on vignettes that included a Black student name after controlling for hours of diversity training. All eight assumptions required were checked using SPSS (Laerd Statistics, n.d.). This study was designed such that the first two assumptions were automatically met, including that there be one continuous dependent variable and at least two independent variables. Although discipline severity was inherently categorical, it was considered to be a continuous variable during data analysis by way of creating a seven-point Likert scale under the assumption that ordinal variables with five or more categories may be treated as continuous (Johnson & Creech, 1983; Norman, 2010; Sullivan & Artino, 2013). The remaining six assumptions were checked, starting with a Durbin-Watson test to ensure independence of residuals. A check for linearity of predictor variables (hours of diversity training and D scores) was completed using scatterplots and partial regression plots. The homoscedasticity of residuals was checked using a residuals plot. To ensure neither independent variable was too highly correlated, a check for multicollinearity was conducted. A check for significant outliers/highly influential points and normal distribution of residuals was completed using a histogram, P-P plot, and Q-Q plot. No data transformations or alternative analyses were required prior to running the regression.

Q2 Do differences exist between the severity of consequences assigned to Black and White student vignettes?

To investigate possible differences in discipline severity between Black and White students, a one-tailed independent samples *t*-test was conducted. The six assumptions required for analysis were met for both groups (Laerd Statistics, n.d.). These included having a continuous dependent variable (discipline rating) and an independent variable consisting of two categorical, independent groups (student names). Additionally, independence of observations, tests for significant outliers, normal distribution of each independent group, and homogeneity of variances were each checked prior to analysis.

Q3 Do differences exist between the severity of consequences assigned to Black and White student vignettes accounting for hours of diversity training?

Three hierarchical multiple regressions were conducted for each vignette to explore potential differences in discipline severity of Black versus White student name after accounting for hours of diversity training. As with the regression for the first research question, all eight assumptions required were checked using SPSS prior to running each of the regressions (Laerd Statistics, n.d.).

### **Summary**

To answer all three research questions, participants were administered a three-part survey that included demographic information, discipline vignettes, and a race IAT using Qualtrics. Prior to collecting data, IRB approval was attained, and amendments were conducted as needed. To analyze participant responses, two statistical methods were used, including hierarchical multiple regression and a one-sided independent samples *t*-test.



## CHAPTER IV

### RESULTS

This study explored the disciplinary decision-making of principals using vignettes and a race IAT. Each participant completed a brief survey that included reporting on basic demographic information about themselves, their role and years of experience at their current school, and the amount of diversity training they have completed in the past three years. They then responded to three discipline vignettes where the name of a student had been manipulated to reflect a common Black name, White name, or name that was considered common across races. Last, participants completed a measure of implicit bias. In doing so, this study aimed to explore whether higher amounts of diversity training could offset the impact of negative racial biases in the discipline process.

#### **Description of Participants**

This study included the responses of 166 high school principals currently working in public schools in the Midwest. Recruitment consisted of contacting high school principals directly via email. Their email addresses were obtained through the departments of education of Wisconsin, Illinois, Michigan, and Indiana. Inclusion criteria consisted of participants being high school principals, assistant principals, vice principals, or deans responsible for disciplinary decisions in their current schools. Principals working in elementary, middle, or private schools were excluded from the participant pools. Additionally, any participant that was not on a computer with a QWERTY keyboard was also excluded to ensure the interactive portion of the survey was completed correctly. To control for this, if-then conditions were created in Qualtrics.

Prior to inviting principals to participate in the study, all of their email addresses were alphabetized and divided into three approximately equal groups to create randomization of participants. A total of 2,715 people were then invited to participate via an initial email. A follow up email was sent one week after to provide participants with a reminder to complete the survey if they had not already done so. Of those invited, 6% completed all but the IAT and 3.5% completed the survey in its entirety. Survey group one consisted of 61 individuals, survey group two consisted of 57, and survey group three consisted of 48. It was decided that the surveys from participants who had not completed the IAT but did complete all three vignettes were kept. This decision was made because the second and third research questions did not require the completion of the IAT and would allow for more power in those analyses. Additionally, sufficient power was reached for the first research question with 97 participants.

Generally, the sample for this study was considered representative of the greater population of high school principals, although principals who identified themselves as White were somewhat overrepresented. For example, 89.2% of the sample was White, 7.8% Black, and 3.0% Hispanic/Latinx. Based on the 2022 national teacher and principal survey conducted by Taie and Lewis (2022) through the Institute of Education Sciences (IES), it was estimated that 77.4% of principals identified as White, 10.6% as Black, and 8.8% as Hispanic. The relatively higher percentage of White principals may be reflective of the narrow geographic region where participants were recruited. In this sample, no participants identified as Asian, Native American/Alaskan Native, or Native Hawaiian/Pacific Islander. Regarding gender, in the IES survey, 64.5% of participants identified as male and 35.5% as female, which was very similar to the 60% and 36.1% of male and female respondents, respectively, in this study. Additionally, in the IES survey, principals had spent on average 4.2 years at their current school with 43.8%

having spent three years or less, 45.2% between 3-9 years, and 11.0% having spent 10 years or more. The only demographic statistics for this study that notably deviated from the national population was years spent at current school, which was notably higher than the national average for this study. However, there was still a reasonable degree of variability across the remaining lesser years of experience categories. Regarding job title, the vast majority of participants listed themselves as “principal”, with the remaining participants having listed either assistant principal or dean. Regarding hours of diversity training, over half of the sample engaged in 10 or more hours of diversity training with the remainder having been sufficiently spread across the other categories of hours offered (i.e., 0-3, 4-6, 7-9). No comparative national surveys were found pertaining to this metric, so it was unknown whether the diversity training of this group was similar to the levels of training among a more representative group. Similarly, comparative national data for job title were not found. Nevertheless, the participants surveyed created an adequate sample for the comparison of various hours of diversity training. Table 5 provides a demographic breakdown of participants.

**Table 5***Summary of Participant Demographic Information*

Baseline Characteristic	<i>n</i>	%
Gender		
Male	103	62.0
Female	60	36.1
Prefer not to say	3	1.8
Race		
White/European American	148	89.2
Black/African American	13	7.8
Hispanic/Latinx	5	3.0
Job title		
Principal	136	81.9
Assistant Principal	26	15.7
Dean	4	2.4
Years in current school		
0-3 years	33	19.9
4-6 years	40	24.1
7-9 years	26	15.7
10+ years	67	40.4
Hours of diversity training completed in the last three years		
0-3 hours	26	15.7
4-6 hours	32	19.3
7-9 hours	18	10.8
10-12 hours	28	16.9

Table 5 (continued)

Baseline Characteristic	<i>n</i>	%
13+ hours	62	37.3
Types of school discipline used in the past year		
Restorative practices	158	95.2
Detention	144	86.7
In School Suspension	146	88.0
Out of School Suspension	156	94.0
Expulsion	86	51.8
Corporal Punishment	0	0.0

*Note.* *N* = 166.

In addition to basic demographic information, participants were also asked to report the types of discipline they have used in the past three years in order to gauge the prevalence of various discipline approaches currently being used in high schools. Research has indicated that out of school suspension is the second most common discipline consequence used in the United States (office discipline referrals being the first; Diem & Welton, 2020; Fisher et al., 2020; Gullo & Beachum, 2020b; Ruiz, 2017). Although this study did not ask participants to report the frequency of the listed disciplinary consequences, the results did indicate that suspensions continue to be a common practice with 94% of participants having endorsed out of school suspension and 88% endorsing the use of in school suspension. Interestingly, the inclusionary discipline response option, restorative practices, was the highest endorsed response at 95.2%.

### **Implicit Racial Bias**

Implicit racial bias was determined using the Greenwald et al. (2022) D score formula. For reference, positive numerical values correspond with pro-White preferences and negative

values correspond with pro-Black preferences; the higher the absolute value of a number, the stronger the preference. An overview of qualitative descriptors for the D score categories and participant outcomes is included in Table 6. The majority of the sample (57.8%) indicated a moderate to strong preference for White names, which aligns with previous implicit bias research (Gullo & Beachum, 2020a; Morin, 2015).

**Table 6**

*Summary of Participant Implicit Racial Bias D Scores*

D Measure Qualitative Descriptor	<i>n</i>	%
Strong Preference for White Names	28	28.9
Moderate Preference for White Names	28	28.9
Slight Preference for White Names	10	10.3
No Preference	23	23.7
Slight Preference for Black Names	3	3.1
Moderate Preference for Black Names	5	5.2
Strong Preference for Black Names	0	0.0

*Note.* *N* = 97.

### **Discipline Severity**

Subjective discipline vignettes were used to investigate possible discipline severity discrepancies between White and Black students. To do so, vignettes of increasing behavior severity were administered, and participants were asked to choose one consequence from a menu of discipline options. To contrive randomization, each student name (Brian, Darius, Jacob) was included in a different vignette for each of the three survey groups. For instance, survey group one received Brian in vignette one, Darius in vignette two, and Jacob in vignette 3 whereas group two received names in the order of Jacob, Brian, and Darius. A general overview of disciplinary choices is included in Table 7 for all participants regardless of which name was included in vignettes. Across all vignettes, the majority of participants chose a non-exclusionary discipline option.

**Table 7***Discipline Consequence Frequencies Assigned by Vignette*

Consequence Assigned	<i>n</i>	%
Vignette 1		
No consequence	29	17.5
Parent phone call	103	62.0
Detention	27	16.3
Refer for tier 2 and behavior support plan	7	4.2
In school suspension	0	0.0
Out of school suspension	0	0.0
Recommend for expulsion	0	0.0
Vignette 2		
No consequence	15	9.0
Parent phone call	123	74.1
Detention	9	5.4
Refer for tier 2 and behavior support plan	17	10.2
In school suspension	1	.6
Out of school suspension	1	.6
Recommend for expulsion	0	0.0
Vignette 3		
No consequence	0	0
Parent phone call	23	13.9
Detention	44	26.5
Refer for tier 2 and behavior support plan	69	41.6
In school suspension	28	16.9
Out of school suspension	2	1.2
Recommend for expulsion	0	0.0

*Note.* *N* = 166. Frequencies reflect discipline consequences regardless of which name was used in vignettes.



### Research Question 1

Q1 After accounting for hours of diversity training, is principal implicit bias a significant predictor of discipline severity on Black student vignettes?

To examine whether principal implicit racial bias predicted the discipline severity of Black students, a hierarchical multiple regression was chosen. Prior to conducting any analyses, a post-hoc power analysis was first completed using G\*Power 3.1.9.7 to support the appropriateness of a regression with a sample size of 97. Using a medium effect size of 0.15 and alpha of .05, the power of the sample size was determined to be .80 ( $F(4,92) = 2.48$ ) and thus sufficient. Second, all necessary assumptions were then assessed and met. The first of these included independence of residuals, which used a Durbin-Watson statistic. Acceptable values are considered to fall between 1.50 and 2.50. The Durbin-Watson statistic for this data was 1.856. Second, linearity of predictor variables, homoscedasticity, the absence of highly influential points, and normal distribution of continuous variables were determined using residual and P-P plots. Third, three dummy variables were created for the vignettes. Because regression models require numerical input to analyze predictions, categorical variables are considered incompatible data. Thus, dummy variables were created by assigning binary numerical placeholders (i.e., 0 and 1) to all three vignettes, allowing them to be included in the regression models. After creating these, a check for multicollinearity was conducted via coefficient correlations. Discipline severity was significantly correlated with all three vignettes, although correlation values were relatively low. Additionally, variance inflation factor (VIF) values were used to determine the effect of possible multicollinearity. A VIF value was produced for each variable within *each* of the regression models created to assess the potential impact of multicollinearity. An acceptable VIF value falls between one and three. Across all models, the VIF values for all variables fell between 1.000 and 1.525, indicating that the models were tolerant of the observed

correlations. Additionally, all three vignettes were significantly correlated with each other. However, because they represented mutually exclusive categories, the presence of one automatically indicates the absence of the others. Said another way, one would be able to consistently predict the presence the outcome of one vignette (which would be the absence the other vignettes) based on the dummy coding of the other vignettes. Thus, the correlation between vignettes was not considered a threat to the overall regression either. Table 8 depicts all coefficient correlations.

After checking all assumptions and creating dummy variables, four models were created to produce the hierarchical regression to explore whether implicit bias scores were predictive of the discipline severity of the name Darius. It was decided that, in order to sufficiently account for the impact of each individual variable in a way that would allow for exploration of the impact of implicit racial bias on discipline severity, four models would be created. The variability that each factor accounted for was determined by the  $R^2$  change after the creation of each model. Vignettes were used to create the initial models so as to account for variations in behavior severity, and thus discipline severity which was expected to increase regardless of student race. Next, hours of diversity training were added to account for any differences in training across participants that may possibly impact discipline decision making. Last, D scores were added to the model to see what remaining variability in discipline severity could be explained by implicit racial bias. Vignettes 1 and 2 produced an  $R^2$  change = .087, the addition of vignette 3 produced a change of .485, and the addition of all remaining variables produced a combined change of .001. Taken together, it was behavior severity that was most predictive of how the name Darius was disciplined, given that the majority of variance (57.4%) was accounted for by the vignettes.

**Table 8**

*Correlations for Discipline Rating, Vignette Number, Hours of Diversity Training, and D Score*

Variable	Vignette 1		Vignette 2		Vignette 3		D Score		Hours of Training	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Darius										
Discipline Rating	-.391	< .001**	-.295	.002**	.754	< .001**	-.002	.491	-.081	.216
Vignette 1			-.391	.000**	-.449	.000**	.005	.481	.043	.339
Vignette 2					-.460	.000**	-.013	.450	.009	.466
Vignette 3							.009	.460	-.057	.292
D Score									-.023	.412

*Note.* Correlation coefficients of all variables and their significance levels.

\* indicates significance at the .05 level, \*\* indicates significance at the .001 level.

Given that the addition of D scores resulted in a change of .000, implicit racial bias was not considered a significant predictor of the discipline severity given to Darius, regardless of hours of diversity training. Table 9 depicts all models included in the hierarchical regression.

### **Research Question 2**

**Q2** Do differences exist between the severity of consequences assigned to Black and White student vignettes?

To examine whether there were differences between the severity of discipline given to vignettes that were assigned a stereotypical Black (i.e., Darius) or White (i.e., Brian) student name, a one-sided independent-samples *t*-test was conducted. The vignette scores for Jacob were removed because that name was considered common across races (Sisense, n.d.; Social Security Administration, n.d.). To compare the discipline of Darius and Brian, all survey groups and vignettes were combined. A post-hoc power analysis was completed to determine the power achieved with the current sample using G\*Power 3.1.9.7 to support the appropriateness of the chosen test. For a one-sided *t*-test with 166 participants, the power achieved was 0.99 and thus considered sufficient. Additionally, prior to completing the *t*-test, all necessary assumptions were tested and met including normality of discipline severity data using Q-Q plots and a test for homogeneity of variance using Levene's test. Surprisingly, the results (see Table 10) indicated no significant difference between the discipline severity between the Black and White student name, which was in contrast to existing research.

**Table 9***Hierarchical Multiple Regression Exploring the Discipline Severity of the Name Darius*

Predictor	<i>B</i>	95% CI for <i>b</i>		<i>SE B</i>	<i>R</i> <sup>2</sup>	<i>R</i> <sup>2</sup> Change
		Lower Limit	Upper Limit			
Model 1					.087	
Vignette 1 (Constant)	2.850**	2.539	3.161	.157		
Vignette 2	-.767*	-1.275	-.259	.256		
<i>R</i> <sup>2</sup> Change Model 1						.087
Model 2					.572	
Vignette 1 (Constant)	1.914**	1.634	2.195	.141		
Vignette 2	.169	-.225	.563	.198		
Vignette 3	2.246**	1.811	2.680	.219		
<i>R</i> <sup>2</sup> Change Model 2						.485
Model 3					.574	
Vignette 1 (Constant)	1.925**	1.542	2.501	.164		
Vignette 2	.169	-.228	.562	.199		
Vignette 3	2.246**	1.802	2.675	.220		
Hours of diversity training	-.030	-.141	.080	.056		
<i>R</i> <sup>2</sup> Change Model 3						.001

Table 9 (continued)

Predictor	<i>B</i>	95% CI for <i>b</i>		<i>SE B</i>	<i>R</i> <sup>2</sup>	<i>R</i> <sup>2</sup> Change
		Lower Limit	Upper Limit			
Model 4					.574	
Vignette 1 (Constant)	2.033**	1.521	2.546	.258		
Vignette 2	.167	-.231	.546	.200		
Vignette 3	2.239**	1.800	2.678	.221		
Hours of diversity training	-.031	-.142	.081	.056		
D score	-.030	-.459	.399	.216		
<i>R</i> <sup>2</sup> Change <i>Model 4</i>						.000
<i>R</i> <sup>2</sup> Total						.574

*Note.* \* indicates  $p < .05$ , \*\* indicates  $p < .01$ .

**Table 10***Independent Samples t-Test Comparing the Discipline Severity Between Darius and Brian*

	Brian		Darius		<i>F</i> (sig.)	<i>t</i>	<i>p</i>	Cohen's <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
Discipline Severity	2.67	1.08	2.61	1.18	1.08 (.300)	.484	.314	.05

*Note.* *p*-value is representative of a one-tailed test.

### Research Question 3

Q3 Do differences exist between the severity of consequences assigned to Black and White student vignettes accounting for hours of diversity training?

To examine whether differences exist between the severity of consequences assigned to Black and White student vignettes after accounting for hours of diversity training, a hierarchical multiple regression was conducted for each vignette. Prior to conducting any analyses, a post-hoc power analysis was completed using G\*Power 3.1.9.7 to support the appropriateness of hierarchical multiple regressions for each of the vignettes. Vignette one achieved a power of .84, vignette two was .82, vignette three was .75, indicating that the sample sizes were large enough to detect a true effect should one exist. Prior to creating the regression models, two dummy variables were created to ensure the two categorical variables, Brian and Darius, were useable in the regression. All necessary assumptions were then assessed and met for each of the regressions. These included the independence of residuals using the Durbin-Watson statistic (all ranged from 1.60-1.837), linearity of predictor variables and homoscedasticity using scatter plots, the absence of highly influential points determined by residual plots, and normal distribution of continuous variables using P-P plots. Additionally, the absence of multicollinearity was determined based on coefficient correlations as well as through VIFs that all fell below the threshold of three. Similar to research question one, in order to conduct the hierarchical multiple regression, three models

were created to assess the variability accounted for by each variable. Models were created by adding variables in the order of Darius, Brian, and hours of diversity training. Table 11 depicts each model and corresponding  $R^2$  change. Taken together, the results of the vignette regressions indicated that hours of diversity training was not a significant predictor of discipline severity. Hours of diversity training only accounted for 0.4%, 3.5%, and 7.3% variance for vignettes one, two, and three, respectively. Of note, the predictive ability of hours of diversity training was approaching significance when behavior severity was higher. Given that observation, it is possible that a larger sample may have provided the response variability needed to detect more statistically meaningful results.



**Table 11**

*Hierarchical Multiple Regression Predicting Differences in Discipline Severity of Darius and Brian When Accounting for Hours of Diversity Training*

Predictor	<i>B</i>	95% CI for <i>B</i>		<i>SE B</i>	<i>R</i> <sup>2</sup>	<i>R</i> <sup>2</sup> Change
		Lower Limit	Upper Limit			
Vignette 1 Model 1					.000	
(Constant)	2.073**	1.939	2.208	.068		
Darius	-.003	-.233	.227	.116		
<i>R</i> <sup>2</sup> Change Model 1						.000
Vignette 1 Model 2					.003	
(Constant)	2.125**	1.922	2.328	.103		
Darius	-.055	-.331	.221	.140		
Brian	-.092	-.364	.179	.138		
<i>R</i> <sup>2</sup> Change Model 2						.003

Table 11 (continued)

Predictor	<i>B</i>	95% CI for <i>B</i>		<i>SE B</i>	<i>R</i> <sup>2</sup>	<i>R</i> <sup>2</sup> Change
		Lower Limit	Upper Limit			
Vignette 1 Model 3					.004	
(Constant)	2.175**	1.870	2.480	.154		
Darius	-.047	-.326	.231	.141		
Brian	-.087	-.360	.187	.138		
Hours of Diversity Training	-.016	-.089	.057	.037		
<i>R</i> <sup>2</sup> Change Model 3						.001
R <sup>2</sup> Total					.004	
Vignette 2 Model 1					.008	
(Constant)	2.267**	2.110	2.424	.079		
Darius	-.152	-.411	.107	.131		
<i>R</i> <sup>2</sup> Change Model 1						.008

Table 11 (continued)

Predictor	<i>B</i>	95% CI for <i>B</i>		<i>SE B</i>	<i>R</i> <sup>2</sup>	<i>R</i> <sup>2</sup> Change
		Lower Limit	Upper Limit			
Vignette 2 Model 2					.032	
(Constant)	2.123**	1.912	2.334	.107		
Darius	-.008	-.302	.286	.149		
Brian	.315	.003	.627	.158		
<i>R</i> <sup>2</sup> Change Model 2						.024
Vignette 2 Model 3					.035	
(Constant)	2.228**	1.864	2.591	.184		
Darius	-.012	-.306	.282	.149		
Brian	.301	-.014	.616	.160		
Hours of Diversity Training	-.029	-.111	.053	.042		
<i>R</i> <sup>2</sup> Change Model 3						.003
R <sup>2</sup> Total:					.035	

Table 11 (continued)

Predictor	<i>B</i>	95% CI for <i>B</i>		<i>SE B</i>	<i>R</i> <sup>2</sup>	<i>R</i> <sup>2</sup> Change
		Lower Limit	Upper Limit			
Vignette 3 Model 1					.027	
(Constant)	3.551**	3.378	3.723	.087		
Darius	.345	0.24	.666	.162		
<i>R</i> <sup>2</sup> Change Model 1						.027
Vignette 3 Model 2					.027	
(Constant)	3.541**	3.300	3.782	.122		
Darius	.355*	-.008	.717	.184		
Brian	.020	-.326	.367	.175		
<i>R</i> <sup>2</sup> Change Model 2						.000

Table 11 (continued)

Predictor	<i>B</i>	95% CI for <i>B</i>		<i>SE B</i>	<i>R</i> <sup>2</sup>	<i>R</i> <sup>2</sup> Change
		Lower Limit	Upper Limit			
Vignette 3 Model 3					.073	
(Constant)	4.013**	3.610	4.416	.204		
Darius	.309	-.047	.666	.180		
Brian	.039	-.300	.378	.172		
Hours of Diversity Training	-.136**	-.231	-.042	.048		
<i>R</i> <sup>2</sup> Change Model 3						.046
R2 Total					.073	

Note. \* indicates  $p < .05$ , \*\* indicates  $p < .01$ .

### **Summary**

In summary, 97 participants completed the full survey and 166 completed all but the IAT. The first research question was answered with the group of 97 who had completed all items, while the second and third questions were answered with the data from all participants. Based on previous research, it was predicted that participants with a stronger preference for White individuals would more harshly discipline vignettes that included the name Darius. It was also expected that there would be differences in discipline severity between Black and White student names, with vignettes including the name Darius being disciplined more harshly. Lastly, it was hypothesized that individuals with greater hours of diversity training would display less discrepancy in their discipline ratings between the Black and White student names. Overall, the results of statistical analyses supported the null hypothesis of each research question. Implicit bias scores were not predictive of the discipline severity of vignettes including a Black student name. Second, no statistically significant difference was found between the discipline severity of Black and White student names across vignettes. Lastly, hours of diversity training did not provide meaningful predictive value when comparing the difference between the discipline severity of the Black and White student names.

## CHAPTER V

### DISCUSSION

The overrepresentation of Black students in the STPP has continued to persist despite decades of knowledge of the issue and ongoing intervention efforts. One of the primary reasons remediation has been difficult is due to how multifaceted the STPP is (Gregory et al., 2010; Novak, 2019). In any one research study, it is only feasible to examine a select few of the dozens of relevant factors. To help contribute to the overall body of research on potential contributing factors, the purpose of this study was to explore how high school principal's implicit racial biases may influence the severity of consequences given to Black students during subjective discipline situations. More specifically, it investigated the interplay between high school principals' implicit racial biases, the amount of diversity training they have engaged in, and how each of those may help explain disciplinary discrepancies between Black and White students.

#### **Overview of Results**

Research question one was intended to explore whether implicit bias scores could be predictive of the discipline severity assigned to vignettes that included the name Darius. Implicit racial bias scores were not predictive of the severity of discipline assigned to vignettes that included the stereotypical Black name, which may have been for several reasons. First, such a result may simply align with existing research that implicit biases are not reliable predictors of human behavior (Forscher et al., 2019; McIntosh et al., 2021). Supporting this thought, the aggregate IAT data from participants was reflective of the general population of overall pro-white preferences (Galvan & Payne, 2024; Gullo, 2017; Gullo & Beachum, 2020a, 2020b), yet

that same IAT did not seem to be reflective of racially biased “behavior” of participants on vignettes. Additionally, IAT D scores accounted for more variance in the regression models predicting discipline severity for Jacob and Brian than of Darius (albeit the effect was not statistically meaningful), further suggesting that IAT scores may not be reliable predictors of behavior. Second, the type of IAT used for this study (name-based) may have impacted how participants disciplined students. Implicit biases are thought to be context dependent (Cooley & Payne, 2017). As such, it may be the case that name was not an activating factor for school discipline vignettes, specifically. That said, there is some research to suggest that it is Black faces (of males in particular) that more quickly and strongly activate negative feelings in White individuals (Staats, 2014). Thus, it is possible that name-based measures (IAT and vignettes) were not activating enough to capture the impact of implicit racial biases in school discipline for this study despite previous research (e.g., Abel & Burger, 2023; Galvan & Payne, 2024) having shown that discrimination does occur against Black individuals based solely on name.

A similar hypothesis pertaining to name and implicit bias activation was also considered for the second research question. Existing research has consistently shown that Black students are disciplined more harshly than White students (Diem & Welton, 2020; Ruiz, 2017). For that reason, the purpose of research question two was intended to replicate prior research demonstrating that Black students are disciplined more harshly than White students. In doing so, such a result would have substantiated the chosen study design. However, no statistically significant differences were found. Therefore, it was again speculated that using stereotypical names alone may simply not have activated racial stereotypes as intended given that there is a well-established body of research showcasing the disciplinary differences between Black and White students. Supporting that hypothesis, a vignette study by Starcke and Porter (2019) was



conducted in which the researchers assessed the disciplinary responses to college student infractions using only common racial names. The researchers similarly did not find disciplinary differences between Black and White students. However, in another vignette study by Taylor and Bailey (2022), the researchers used common Black and White names *in addition to* racial descriptives and a student photograph. In their study, they detected differences in how participants conceptualized vignettes based on race. The latter two studies, taken together in combination with the current study, suggest the possible need to include more than a name when conducting vignette research related to implicit racial biases. Regarding both research questions one and two, it is also possible that the intent of the study was easily discernable to participants, thus leading to more socially desirable responding.

Regarding research question three, to date (to the knowledge of this researcher), no study has yet examined how the quantity of diversity training hours a principal has engaged in may influence how they discipline the subjective behaviors of Black versus White students. For this reason, the third research question was intended to help determine if discipline discrepancies existed when holding hours of diversity training constant. The purpose of doing so was to explore the possible bias-mitigating potential of engaging in higher amounts of diversity training. Results indicated that hours of diversity training were not a statistically meaningful predictor of the severity of discipline assigned to the vignettes including the Black student name. Of importance though, the variance explained by the addition of hours of diversity training into the regression models increased as the behavior severity of vignettes increased. Thus, although the predictive power of hours of diversity training was not statistically meaningful in this study, it began approaching significance. Future studies that achieve larger sample sizes may explore the

potential mitigating effects of diversity training on disciplinary decisions of more severe subjective behaviors.

### **Implications**

This study was intended to help better understand the role of school principals in exclusionary discipline practices that contribute to the STPP. As the disciplinary gatekeepers, and the leaders of their schools, they should not be overlooked when studying the STPP. Implicit racial biases have been shown to impact discipline decision making (Gullo, 2017; Gullo & Beachum, 2020a, 2020b; Jarvis & Okonofua, 2020); however, limited research has examined implicit racial bias specifically with school principals and their discipline decision making. Additionally, research on the potential mitigating power of diversity training during subjective discipline situations with Black students is lacking. The results of this study indicated that implicit racial biases were not predictive of how Black students were disciplined. Additionally, having engaged in more hours of diversity training was not strongly predictive of lesser consequences for Black students either. Taken together, the results of this study suggest that principals' implicit racial biases and the hours of diversity training they have engaged in do not help predict or explain discipline discrepancies between Black and White students within the structure of this particular study. For these reasons, research is still needed to identify areas to intervene in the racial discipline gap as it pertains to subjective disciplinary situations. One possible avenue for future research related to implicit bias and subjective discipline may be guided by a recent study by Markowitz et al. (2023). The authors focused on the possible interplay between principals' discipline decisions and teachers' language use in discipline referrals. For instance, when evaluating teachers' language patterns in discipline referrals, those given to Black students contained more emotionally negative content than those for White

students. Additionally, the details included in the referral and description of events was linguistically different for Black students than White students. Considering the Markowitz et al. (2023) findings, future research on subjective discipline, principals, and the racial discipline gap might focus on how teachers' implicit biases might produce referrals that are linguistically and affectively different for Black students, thereby inadvertently influencing principals' decision making or activating their biases.

It could also be possible that the results of this study reflect legitimate social change and/or the impact of current events. This study was conducted amidst the resurgence of the Black Lives Matter and Antiracism movements that were fueled by the deaths of George Floyd and Breonna Tylor, the 2020 presidential elections, a global pandemic, and the subsequent contentions surrounding race-related policy in schools. As a result, the national political climate surrounding race was highly charged when this survey was administered. It is possible that socially desirable responding was stronger than anticipated during this study. That is to say, within the context of current events, responses may have been reflective of social desirability that permeates daily life, not just the survey setting. Marcucci (2020) stated, "In the current American era, accusations of racism are deeply offensive to many" (p. 50), arguably more so than in decades past. The author also discussed how the heightened awareness of racial biases could lead to behavior change that is fueled by a desire to be perceived as race-neutral and therefore socially acceptable. Taking Marcucci's perspective into consideration, social desirability on the survey may be reflective of the greater political context of the country and the subsequent heightened pressure for social desirability in daily interactions. Herbstrith and Busse (2020) similarly speculated about how social desirability might override implicit biases by way of creating "normative pressure". Meaning that, as awareness of racial tensions heightens,

increased expectations for individuals to act in socially desirable ways can result in shifts in behavioral norms--which is to be “socially desirable”, or not racist, in *all* contexts. Although it is not possible to tell if the results of this study are the result of increasing normative pressures on a large scale and/or increased desire to be seen as race neutral in the survey and beyond, the context of the political climate of the nation at the time of this study should not be discounted.

### **Limitations**

There were several limitations important to this study. First, it used a vignette methodology where participants were asked to respond to three vignettes in which the name of the ‘student’ had been manipulated to reflect a stereotypical Black name, White name, or name common across races. To accomplish this, three separate groups of participants were surveyed with the intent of randomizing the presentation of student names. For example, survey group one received Brian in vignette one, Darius in vignette two, and Jacob in vignette three whereas survey group two received the names in the order of Jacob, Darius, and Brian. In doing this, the behavior severity of the vignettes inadvertently became an additional variable, which convoluted statistical analyses. Second, despite efforts to strategically organize the survey such that participants would be less likely to respond in a socially desirable way, it is thought that they were able to quickly determine the general purpose of the study. Multiple participants contacted the researcher to share their opinions about bias research (and their reservations about completing the survey due to it being about racial biases) despite there being no outright disclosure in the informed consent or survey that bias research was being conducted. Furthermore, given that the results of this study did not align with preexisting research showing that Black students are disciplined more harshly than White students, it is thought that name-based vignettes may not have been a viable instrument for discerning implicit racial biases

during subjective discipline. Lastly, Greenwald et al. (2022) reported that the test-retest reliability of one single IAT administration for one individual has an  $r = .50$ . Thus, to best interpret the reliability of one individual's implicit bias, multiple administrations of an IAT is recommended. Without conducting multiple administrations, Greenwald et al. (2022) report that it is the aggregate data, which has an  $\alpha = .80$ , that would be most reflective of the overall implicit biases of a sample. Given that multiple administrations per participant was not feasible in this study, the interpretation of the first research question is reflective of aggregate data, as recommended by Greenwald et al. (2022).

### **Future Research**

Although principals' implicit racial bias scores and hours of diversity training did not help explain differences between Black and White students in this particular study, their role in the racial discipline gap is still a fruitful area of study. One avenue for future research could be focusing more specifically on the various types of diversity training. Although participants in this study were provided with a definition of training to help with standardization, it was impossible to know what they considered diversity training in responding to this question. Further, the quality of those trainings, the similarity of them across participants, and the accuracy of participant's estimations could not be verified. One such possibility for developing a better understanding of diversity training for principals might include a closer examination of their training programs. Emerging research has suggested that integrated forms of diversity training may have more meaningful effects on culturally responsive behavior in comparison to stand alone trainings (Barakat et al., 2021; Forscher et al., 2019; Hall & Theriot, 2016). Additionally, research has also shown that practicing principals may not have received formal training, or enough training, in culturally responsive practices (Marchioli et al., 2020; Minkos et al., 2017;

Moughania, 2018). Thus, a future avenue of research could consist of exploring principals' perception of their diversity knowledge and if those perceptions are predictive of disciplinary decision making.

Another possible avenue for expanding on the results of this study could include broadening the scope of impact of principals' implicit racial biases and diversity training. The reason being is that principals' one-to-one interactions with students may not be as influenced by implicit bias or diversity training as this researcher thought. Rather, how principals' diversity training and/or implicit racial biases impact their approach to, and implementation of, systemic interventions may be of greater utility in exploring their role in the discipline gap. For instance, a study conducted by McIntosh et al. (2020) explored a similar idea by providing 35 K-12 school principals with monthly discipline data depicting Black-White office discipline referral differences for five months. At the end of the intervention, only one school implemented a school improvement plan, suggesting that having the needed data to be informed of discrepancies is not enough to elicit change behavior. Rather, they may possibly need training on actionable ways to intervene in addition to being provided with data. Further driving this point, Ferguson et al. (2023) conducted a study that explored how principal attitudes influenced discipline. They surveyed participants to see how an "exclusionary mindset" impacted discipline in comparison to a "prevention mindset". In their study, an exclusionary mindset meant principals had a propensity for exclusionary consequences, whereas those with a prevention mindset tended towards restorative approaches such as providing students with support (e.g., counseling) or getting them more involved (e.g., extracurricular participation). Their findings indicated that exclusionary mindsets were positively correlated with greater discipline severity and prevention mindsets were inversely correlated with the amount of discipline students received, suggesting

that personal beliefs and perspectives about specific discipline types, and not necessarily bias, likely have a notable impact on broader school systems.

Additionally, the individualization required to adequately address diversity needs can vastly differ from one school system to the next. Therefore, research exploring how principals' implicit biases and/or diversity training experiences impact the effectiveness of systemic interventions geared toward closing the racial discipline gap could also be useful. McIntosh et al. (2021) conducted a recent study that focused on an emerging culturally responsive positive behavior intervention and supports (CR-PBIS) model that is specifically geared

towards individualized intervention. The authors studied the potential effectiveness of ReACT (Racial equity through Assessing data for vulnerable decision points, Culturally responsive behavior strategies, and Teaching about implicit bias and how to neutralize it), which was designed to be a school-wide, integrated professional development system. Meaning, it is meant to consist of diversity training that is ongoing and embedded across all aspects of staff's job (e.g., meetings, supervision, classroom instruction, research, data disaggregation and interpretation, discipline, etc.), to better allow for individualization of both student interventions and staff training that meets the unique needs of an individual school system. The guideposts for such individualization within a PBIS framework included highly specific root cause analyses of ODRs (e.g., when and where behaviors occur each day), identifying vulnerable decision points (times when bias is likely to be problematic) and creating well-defined bias neutralizing routines for these, and using personal matrices to bridge expectations at school and home to allow for community individualization (McIntosh et al., 2021). Taken together, principals are the guiding force behind a model like ReACT and other CR-PBIS approaches. Their personal beliefs and diversity training experiences likely impact how they train/support their staff, individualize

systemic interventions, and interpret discipline data though, which helps affirm the need for continued research on their role in the discipline gap.

### **Conclusion**

This study was conducted in hopes of contributing to the ongoing efforts to mitigate the overrepresentation of Black students in exclusionary discipline and subsequently the STPP. Because school principals' potential role in exacerbating and mitigating the STPP has been minimally studied, this study was conducted as a way to further investigate whether their implicit biases and diversity training experiences would help explain disciplinary discrepancies between Black and White students during subjective discipline. Results indicated that neither implicit racial biases nor greater quantities of diversity training helped to explain principals' roles in the racial discipline gap. Given that the results of this study did not align with existing research on discipline discrepancies between Black and White students, it is thought that using a name-based design may have impeded meaningful findings. Despite statistically insignificant findings, the racial discipline gap and subsequent STPP are still an important area of study. More specifically, principals' roles in the discipline gap, and ways they can intervene, also continues to be a much-needed area of study. As such, it is hoped that the results of this study will help inform future implicit bias research with principals as well as future implicit bias research that utilizes a vignette design.



## REFERENCES

- Abel, M., & Burger, R. (2023). Unpacking name-based race discrimination (Discussion Paper Series no. 16245). *Institute of Labor Economics*.  
[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4490163](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4490163)
- Annie E. Casey Foundation. (2021). *Youth incarceration in the United States*.  
<https://assets.aecf.org/m/resourcedoc/aecf-youthIncarcerationinfographic-2021.pdf>
- Bal, A., Afacan, K., & Cakir, H. I. (2018). Culturally responsive school discipline: Implementing learning lab at a high school for systemic transformation. *American Educational Research Journal*, 55(5), 1007-1050. <https://doi.org/10.3102/0002831218768796>
- Barakat, M., Reyes-Guerra, D., Stefanovic, M., & Shatara, L. (2021). An examination of the relationship between a redesigned school leadership preparation program and graduates' cultural competence. *Leadership and Policy in Schools*, 20(1), 21-40.  
<https://doi.org/10.1080/15700763.2020.1833937>
- Barnes, J., & Motz, R. (2018). Reducing racial inequalities in adulthood arrest by reducing inequalities in school discipline: Evidence from the school-to-prison pipeline. *Developmental Psychology*, 54(12), 2328-2340. <https://doi.org/10.1037/dev0000613>
- Bezrukova, K., Jehn, K. A., & Spell, C. S. (2012). Reviewing diversity training: Where we have been and where we should go. *Academy of Management Learning & Education*, 11(2), 207-227. <https://doi.org/10.5465/amle.2008.0090>

- Bezrukova, K., Spell, C. S., Perry, J. L., & Jehn, K. A. (2016). A meta-analytical integration of over 40 years of research on diversity training evaluation. *Psychological Bulletin*, 142(11), 1227-1274. <https://doi.org/10.1037/bul0000067>
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.
- Brown, A. (2018). From subhuman to human kind: Implicit bias, racial memory, and black males in schools and society. *Peabody Journal of Education*, 93(1), 52-65. <https://doi.org/10.1080/0161956X.2017.1403176>
- Bursell, M., & Olsson, F. (2021). Do we need dual-process theory to understand implicit bias? A study of the nature of implicit bias against Muslims. *Poetics (Amsterdam)*, 87, 101549. <https://doi.org/10.1016/j.poetic.2021.101549>
- Cagle, J. F. (2017). The cost of color in public education--an examination of disproportionate suspensions. *Journal of Organizational and Educational Leadership*, 3(1), 1-33. <http://files.eric.ed.gov/fulltext/EJ1161987.pdf>
- Carpenter, T., Pogacar, R., Pullig, C., Kouril, M., Aguilar, S., LaBouff, J. P., Isenberg, N., & Chakroff, A. (2019). Survey-software implicit association tests: A methodological and empirical analysis. *Behavior Research Methods*, 51(5), 2194-2208. <https://doi.org/10.3758/s13428-019-01293-3>
- Carter, P., Skiba, R., Arredondo, M. I., & Pollock, M. (2017). You can't fix what you don't look at: Acknowledging race in addressing racial discipline disparities. *Urban Education*, 52(2), 207-235. <https://doi.org/10.1177/0042085916660350>
- Cooley, E., & Payne, B. K. (2017). Using groups to measure intergroup prejudice. *Personality & Social Psychology Bulletin*, 43(1), 46-59. <https://doi.org/10.1177/0146167216675331>

- Cox, A. W. (2017). *Preparing education leaders to meet diverse school needs: A cross-case analysis of principal preparation programs* (Publication No. 10685569). [Doctoral dissertation, Clemson University]. ProQuest Dissertations & Theses Global.
- DeMatthews, D. (2015). Making sense of social justice leadership: A case study of a principal's experiences to create a more inclusive school. *Leadership in Policy in Schools* 14(2), 139-166. <https://doi.org/10.1080/15700763.2014.997939>
- DeMatthews, D. (2016a). Effective leadership is not enough: Critical approaches to closing the racial discipline gap. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 89(1), 7-13. <https://doi.org/10.1080/00098655.2015.1121120>
- DeMatthews, D. (2016b). The racial discipline gap: Critically examining policy, culture, and leadership in a struggling urban district. *Journal of Cases in Educational Leadership*, 19(2), 82-96. <https://doi.org/10.1177/1555458915626758>
- DeMatthews, D., Carey, R., Olivarez, A., & Moussavi Saeedi, K. (2017). Guilty as charged? Principals' perspectives on disciplinary practices and the racial discipline gap. *Educational Administration Quarterly*, 53(4), 519-555. <https://doi.org/10.1177/0013161X17714844>
- Diem, S., & Welton, A. D. (2020). *Anti-racist educational leadership and policy: Addressing racism in public education*. Taylor and Francis. <https://doi.org/10.4324/9780429487224>
- Every Student Succeeds Act, Pub. L. No. 114-95, 129 Stat. 1802 (2015). <https://www.congress.gov/bill/114th-congress/senate-bill/1177>
- Farley, A., Childs, J., & Johnson, O. (2019). Preparing school leaders for America's wicked problems? How the revised PSEL and NELP standards address equity and justice. *Education Policy Analysis Archives* 27(115), 1-30. <https://doi.org/10.14507/epaa.27.4229>

- Fedor, C. G. (2014). Stereotypes and prejudice in the perception of the “other”. *Procedia - Social and Behavioral Science*, 149, 321-326. <https://doi.org/10.1016/j.sbspro.2014.08.257>
- Fenning, P., Golomb, S., Gordon, V., Kelly, M., Scheinfeld, R., & Morello, T. (2008). Written discipline policies used by administrators: Do we have sufficient tools of the trade? *Journal of School Violence*, 7(2), 123-146. [https://doi.org/10.1300/J202v07n02\\_08](https://doi.org/10.1300/J202v07n02_08)
- Fenning, P., Pulaski, S., Gomez, M., Morello, M., Maciel, L., Maroney, E., & Maltese, R. (2012). Call to action: A critical need for designing alternatives to suspension and expulsion. *Journal of School Violence*, 11(2), 105-117. <https://doi.org/10.1080/15388220.2011.646643>
- Fenning, P., & Rose, J. (2007). Overrepresentation of African American students in exclusionary discipline: The role of school policy. *Urban Education* 42(6), 536-559. <https://doi.org/10.1177/0042085907305039>
- Ferguson, Z. E., Jarvis, S. N., Antonoplis, S., & Okanofua, J. A. (2023). Principal beliefs predict responses to individual students’ misbehavior. *Educational Researcher*, 52(5), 315-319. <https://doi.org/10.3102/0013189X231158389>
- Fisher, B., Dawson-Edwards, C., & Higgins, E., & Swartz, K. (2020). Who belongs in school? Examining the link between Black and White racial disparities in sense of school belonging and suspension. *Journal of Community Psychology* 48(5), 1481-1499. <https://doi.org/10.1002/jcop.22342>
- Fiske, S. T., & Taylor, S. E. (2008). *Social cognition: From brains to culture* (1st ed.). McGraw-Hill Higher Education.

Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B.

A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*, 117(3), 522-559.

<https://doi.org/10.1037/pspa0000160>

Galvan, M. J., & Payne, B. K. (2024). Implicit bias as a cognitive manifestation of systemic racism. *Daedalus*, 153(1), 106. [https://doi.org/10.1162/daed\\_a\\_02051](https://doi.org/10.1162/daed_a_02051)

Girvan, E., Gion, C., McIntosh, K., & Smolkowski, K. (2016). The relative contribution of subjective office referrals to racial disproportionality in school discipline. *School Psychology Quarterly*, 32(3), 392-404. <https://doi.org/10.1037/spq0000178>

Graves, S., & Howes, C. (2011). Ethnic differences in social-emotional development in preschool: The impact of teacher child relationships and classroom quality. *School Psychology Quarterly* 26(3), 202-214. <https://doi.org/10.1037/a0024117>

Greenwald, A. G., Brendl, M., Cai, H., Cvencek, D., Dovidio, J. F., Friesse, M., Hahn, A.,  
Hehman, E., Hofmann, W., Hughes, S., Hussey, I., Jordan, C., Kirby, T. A., Lai, C. K.,  
Lang, J. W. B., Lindgren, K. P., Maison, D., Ostafin, B. D., Rae, J. R., . . . Wiers, R. W.  
(2022). Best research practices for using the implicit association test. *Behavior Research Methods*, 54(3), 1161-1180. <https://doi.org/10.3758/s13428-021-01624-3>

Greenwald, A. G., & Krieger, L. H. (2006). Implicit bias: Scientific foundations. *California Law Review*, 94(4), 945-967. <https://doi.org/10.2307/20439056>

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464-1480. <https://doi.org/10.1037/0022-3514.74.6.1464>

- Gregory, A., Skiba, R., & Mediratta, K. (2017). Eliminating disparities in school discipline: A framework for intervention. *Review of Research in Education, 41*, 253-278.  
<https://doi.org/10.3102/0091732X17690499>
- Gregory, A., Skiba, R., & Noguera, P. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher 39*(1), 59-68.  
<https://doi.org/10.3102/0013189X09357621>
- Gullo, G. (2017). *Implicit bias in school disciplinary decisions* (Publication No. 10278566). [Doctoral dissertation, Leigh University]. ProQuest & Theses Global.
- Gullo, G. (2020). Equity endeavors through the justice for bias framework. In P. Youngs, J. Kim, & M. Mavrogordato (Eds.), *Exploring principal development and teacher outcomes: How principals can strengthen instruction, teacher retention, and student achievement* (pp. 86-100). Routledge. <https://doi.org/10.4324/9780429356247>
- Gullo, G., & Beachum, F. (2020a). Does implicit bias matter at the administrative level? A study of principal implicit bias and the racial discipline severity gap. *Teachers College Record: The Voice of Scholarship in Education, 122*(3), 1-28.  
<https://doi.org/10.1177/016146812012200309>
- Gullo, G., & Beachum, F. (2020b). Principals navigating discipline decisions for social justice: An informed grounded theory study. *Heliyon, 6*(12), 1-11.  
<https://doi.org/10.1016/j.heliyon.2020.e05736>
- Gullo, G., Capatosto, C., & Staats, C. (2018). *Implicit bias in schools: A practitioner's guide*. Taylor and Francis. <https://doi.org/10.4324/9781351019903>

- Haider, A. H., Sexton, J., Sriram, N., Cooper, L. A., Efron, D. T., Swoboda, S., Villegas, C. V., Haut, E. R., Bonds, M., Pronovost, P. J., Lipsett, P. A., Freischlag, J. A., & Cornwell, E. (2011). Association of unconscious race and social class bias with vignette-based clinical assessments by medical students. *JAMA: The Journal of the American Medical Association*, 306(9), 942-951. <https://doi.org/10.1001/jama.2011.1248>
- Hall, J. C., & Theriot, M. T. (2016). Developing multicultural awareness, knowledge, and skills: Diversity training makes a difference? *Multicultural Perspectives*, 18(1), 35-41. <https://doi.org/10.1080/15210960.2016.1125742>
- Harper, K. (2020). School discipline reform is still needed, but is discipline policy still the solution? *National Association of State Boards of Education*, 20(2), 43-47. [https://nasbe.nyc3.digitaloceanspaces.com/2020/05/Harper\\_May-2020-Standard.pdf](https://nasbe.nyc3.digitaloceanspaces.com/2020/05/Harper_May-2020-Standard.pdf)
- Hartley, J. (1998). *Learning and studying: A research perspective*. Routledge. <https://doi.org/10.4324/9780203132951>
- Hemez, P., Brent, J. J., & Mowen, T. J. (2020). Exploring the school-to-prison pipeline: How school suspensions influence incarceration during young adulthood. *Youth Violence and Juvenile Justice*, 18(3), 235-255. <https://doi.org/10.1177/1541204019880945>
- Herbsttrith, J. C., & Busse, G. A. (2020). Seven million and counting: Systems-level consultation for students of LGBTQ+ families. *Journal of Educational and Psychological Consultation*, 30(1), 29-62. <https://doi.org/10.1080/10474412.2019.1565539>
- Hockenberry, S. (June 2020). *Juvenile justice statistics national report series bulletin*. U.S. Department of Justice. <https://ojjdp.ojp.gov/sites/g/files/xyckuh176/files/media/document/juveniles-in-residential-placement-2017.pdf>

- Hussain, A., Garavan, T., Carbery, R., O'Brien, F., & McGuire, D. (2016). Diversity training programme outcomes: A systemic review. *Human Resource Development Quarterly*, 27(1). <https://doi.org/10.1002/hrdq.21221>
- IATGEN. (n.d.). *About the IAT*. <https://iatgen.wordpress.com/about-iat/>
- Jackson, S. M., Hillard, A. L., & Schneider, T. R. (2014). Using implicit bias training to improve attitudes toward women in STEM. *Social Psychology of Education*, 17(3), 419-438. <https://doi.org/10.1007/s11218-014-9259-5f>
- Jarvis, S. N., & Okonofua, J. A. (2020). School deferred: When bias affects school leaders. *Social Psychological & Personality Science*, 11(4), 492-498. <https://doi.org/10.1177/1948550619875150>
- Johnson, D. R., & Creech, J. R. (1983). Ordinal measures in multiple indicator models: A simulation study of categorization error. *American Sociological Review*, 48, 398-407.
- Keleher, T. (2000). *Racial disparities related to school zero tolerance policies: Testimony to the U.S. commission on civil rights* (Report No. ED454324). Applied Research Center. <http://files.eric.ed.gov/fulltext/ED454324.pdf>
- Khalifa, M. A., Gooden, M. A., & Davis, J. E. (2016). Culturally responsive school leadership: A synthesis of the literature. *Review of Educational Research*, 86(4), 1272-1311. <https://doi.org/10.3102/0034654316630383>
- Kurdi, B., Carroll, T. J., & Banaji, M. R. (2021). Specificity and incremental predictive validity of implicit attitudes: Studies of a race-based phenotype. *Cognitive Research: Principles and Implications*, 6(1), 61-61. <https://doi.org/10.1186/s41235-021-00324-y>



- Kurdi, B., Ratliff, K. A., & Cunningham, W. A. (2021). Can the implicit association test serve as a valid measure of automatic cognition? A response to Schimmack (2021). *Perspectives on Psychological Science*, 16(2), 422-434. <https://doi.org/10.1177/1745691620904080>
- Laerd Statistics. (n.d.). *Multiple regression analysis using SPSS statistics*.  
<https://statistics.laerd.com/spss-tutorials/multiple-regression-using-spss-statistics.php>
- Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y., J., Pedram, C., Marshburn, C. K., Simon, S., Blanchar, J. C., Joy-Gaba, J. A., Conway, J., Redford, L., Klein, R. A., Roussos, G., Schellhaas, F. M. H., Burns, M., . . . Nosek, B. A. (2016). Reducing implicit racial preferences: II. intervention effectiveness across time. *Journal of Experimental Psychology: General*, 145(8), 1001-1016.  
<https://doi.org/10.1037/xge0000179>
- Law Library. (n.d.). *Zero tolerance*. Law Library-American Law and Legal Information.  
<https://law.jrank.org/pages/11439/Zero-Tolerance.html>
- Lindsey, A., King, E., Hebl, M., & Levine, N. (2014). The impact of method, motivation, and empathy on diversity training effectiveness. *Journal of Business and Psychology*, 30(3), 605-617. <https://doi.org/10.1007/s10869-014-9384-3>
- Llorente, M. (2014). Help us dismantle the school-to-prison pipeline. *Children's Rights Litigation*, 16(3), 13-21.  
[https://www.americanbar.org/content/dam/aba/publications/litigation\\_committees/childrights/spring2014.pdf](https://www.americanbar.org/content/dam/aba/publications/litigation_committees/childrights/spring2014.pdf)

- Marchioli, K., Gray, C., & Wright, C. (2020). *Principal preparation programs for change-based leadership: Reimagining principal preparation programs at riverside university* (Publication No. 27960978). [Doctoral dissertation, University at Buffalo]. ProQuest & Theses Global.
- Marcucci, O. (2020). Implicit bias in the era of social desirability: Understanding antiblackness in rehabilitative and punitive school discipline. *The Urban Review*, 52(1), 47-74.  
<https://doi.org/10.1007/s11256-019-00512-7>
- Markowitz, D. M., Kittelman, A., Girvan, E. J., Santiago-Rosario, M. R., & McIntosh, K. (2023). Taking note of our biases: How language patterns reveal bias underlying the use of office discipline referrals in exclusionary discipline. *Educational Researcher*, 52(9), 525-534.  
<https://doi.org/10.3102/0013189X231189444>
- McCarthy, M., Shelton, S., & Murphy, J. (2016). Policy penetration of the ISLLC Standards. *Leadership and Policy in Schools* 15(2), 221-230.  
<https://doi.org/10.1080/15700763.2015.1093149>
- McGlone, M. S., & Pfister, R. A. (2007). The generality and consequences of stereotype threat. *Sociology Compass*, 1(1), 174-190. <https://doi.org/10.1111/j.1751-9020.2007.00021.x>
- McIntosh, K., Girvan, E. J., Fairbanks Falcon, S., McDaniel, S. C., Smolkowski, K., Bastable, E., Santiago-Rosario, M. R., Izzard, S., Austin, S. C., Nese, R. N. T., & Baldy, T. S. (2021). Equity-focused PBIS approach reduces racial inequities in school discipline: A randomized controlled trial. *School Psychology*, 36(6), 433-444.

- McIntosh, K., Smolkowski, K., Gion, C. M., Witherspoon, L., Bastable, E., & Girvan, E. J. (2020). Awareness is not enough: A double-blind randomized controlled trial of the effects of providing discipline disproportionality data reports to school administrators. *Educational Researcher*, 49(7), 533-537. <https://doi.org/10.3102/0013189X20939937>
- Metropolitan Council of Educational Administration Programs Committee on Cultural Responsiveness. (2017, November 28). *Cultural responsiveness: Definitions and principles*. <http://www.nysed.gov/common/nysed/files/principal-project-phase-2-mceap-definition-and-guiding-principles-for-cultural-responsiveness-and-cultural-competency.pdf>
- Miller, C. M., & Martin, B. N. (2015). Principal preparedness for leading in demographically changing schools: Where is the social justice training? *Educational Management, Administration & Leadership*, 43(1), 129-151. <https://doi.org/10.1177/1741143213513185>
- Minkos, M. L., Sassu, K. A., Gregory, J. L., Patwa, S. S., Theodore, L. A., & Femc-Bagwell, M. (2017). Culturally responsive practice and the role of school administrators. *Psychology in the Schools*, 54(10), 1260-1266. <https://doi.org/10.1002/pits.22072>
- Morin, R. (2015, August 19). *Exploring racial bias among biracial and single-race adults: The IAT*. Pew Research Center. <https://www.pewresearch.org/social-trends/2015/08/19/exploring-racial-bias-among-biracial-and-single-race-adults-the-iat/>
- Moughania, A. B. (2018). *Intercultural proficiency of principals leading IB schools in urban America: A qualitative study* (Publication No. 10974267). [Doctoral dissertation, Loyola University Chicago]. ProQuest & Theses Global.

Murphy, J. (2005). Unpacking the foundations of ISLLC standards and addressing concerns in the academic community. *Educational Administration Quarterly* 41(1), 154-191.

<https://doi.org/10.1177/0013161X04269580>

National Association of School Psychologists. (2019). *Key messages and talking points for school psychologists: Advancing social justice*. <https://www.nasponline.org/x53289.xml>

National Center for Cultural Competence. (n.d.). *Conscious & unconscious biases in health care*.

Georgetown University for Center for Child & Human Development.

<https://nccc.georgetown.edu/bias/index.php>

National Center for Education Statistics. (2019). *Spotlight A: Characteristics of public school teachers by race/ethnicity*. [https://nces.ed.gov/programs/raceindicators/spotlight\\_a.asp](https://nces.ed.gov/programs/raceindicators/spotlight_a.asp)

National Center for Education Statistics. (2020a). *Characteristics of public school principals*.

<https://nces.ed.gov/programs/coe/indicator/cls>

National Center for Education Statistics. (2020b). *Race and ethnicity of public school teachers and their students*. <https://nces.ed.gov/pubs2020/2020103.pdf>

National Center for Education Statistics. (2022). *Characteristics of 2020–21 Public and Private K–12 School Principals in the United States: Results from the National Teacher and Principal Survey First Look (NCES 2022-112)*.

<https://nces.ed.gov/pubs2022/2022113.pdf>

National Policy Board for Educational Administration. (2011). *Educational leadership program standards: 2011 ELCC building level*. [https://www.npbea.org/wp-](https://www.npbea.org/wp-content/uploads/2019/01/ELCC-Building-Level-Standards-2011.pdf)

[content/uploads/2019/01/ELCC-Building-Level-Standards-2011.pdf](https://www.npbea.org/wp-content/uploads/2019/01/ELCC-Building-Level-Standards-2011.pdf)

National Policy Board for Educational Administration. (2015). Professional standards for educational leaders 2015. <https://www.npbea.org/psel/>

- National Policy Board for Educational Administration. (2018). *National Educational Leadership Preparation (NELP) program standards building level*. <https://www.npbea.org/wp-content/uploads/2018/11/NELP-Building-Standards.pdf>
- Noguera, P. A. (2009). Crossing borders, breaking barriers: How teachers can transcend race, class and cultural differences to promote academic achievement and engagement for all. *Ethnicity and Race in a Changing World*, 1(2), 50-53.
- Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, 15(5), 625-632. <https://doi.org/10.1007/s10459-010-9222-y>
- Novak, A. (2019). The school-to-prison pipeline an examination of the association between suspension and justice system involvement. *Criminal Justice and Behavior*, 46(8), 1165-1180. <https://doi.org/10.1177/0093854819846917>
- 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. <https://doi.org/10.1037/pspa0000023>
- Owens, J., & McLanahan, S. S. (2020). Unpacking the drivers of racial disparities in school suspension and expulsion. *Social Forces*, 98(4), 1548-1577. <https://doi.org/10.1093/sf/soz095>
- Pas, E. T., Bradshaw, C. P., & Mitchell, M. M. (2011). Examining the validity of office discipline referrals as an indicator of student behavior problems. *Psychology in the Schools*, 48(6), 541-555. <https://doi.org/10.1002/pits.20577>

- Pinto, L. E. (2013). *From discipline to culturally responsive engagement: 45 classroom management strategies* (1st ed.). Corwin. <https://doi.org/10.4135/9781452285160>
- Project Implicit. (n.d.). *Frequently asked questions*.  
<https://implicit.harvard.edu/implicit/faqs.html>
- Redfield, S., & Nance, J. P. (2016). The American bar association joint task force on reversing the school-to-prison pipeline preliminary report. *The University of Memphis Law Review*, 47(1), 1-180.
- Rovner, J. (2021). Black disparities in youth incarceration. *The Sentencing Project*.  
<https://www.sentencingproject.org/wp-content/uploads/2017/09/Black-Disparities-in-Youth-Incarceration.pdf>
- Ruhl, C. (2021, May 04). *Cognitive bias examples*. Simply Psychology.  
[www.simplypsychology.org/cognitive-bias.html](http://www.simplypsychology.org/cognitive-bias.html)
- Ruiz, R. (2017). School-to-prison pipeline: An evaluation of zero tolerance policies and their alternatives. *Houston Law Review*, 54(3), 803-838.  
<https://houstonlawreview.org/article/3920-school-to-prison-pipeline-an-evaluation-of-zero-tolerance-policies-and-their-alternatives>
- Sander, J. (2010). School psychology, juvenile justice, and the school to prison pipeline. *Communique*, 39(4), 2-4.
- Sander, J. B., Sharkey, J. D., Groomes, A. N., Krumholz, L., Walker, K., & Hsu, J. Y. (2011). Social justice and juvenile offenders: Examples of fairness, respect, and access in education settings. *Journal of Educational and Psychological Consultation*, 21(4), 309-337. <https://doi.org/10.1080/10474412.2011.620816>

- Sander, J. B., Sharkey, J. D., Olivarri, R., Tanigawa, D. A., & Mauseth, T. (2010). A qualitative study of juvenile offenders, student engagement, and interpersonal relationships: Implications for research directions and preventionist approaches. *Journal of Educational Psychological Consultation*, 20(4), 288-315.  
<https://doi.org/10.1080/10474412.2010.522878>
- Schimmack, U. (2021a). Invalid claims about the validity of implicit association tests by prisoners of the implicit social-cognition paradigm. *Perspectives on Psychological Science*, 16(2), 435-442. <https://doi.org/10.1177/1745691621991860>
- Schimmack, U. (2021b). The implicit association test: A method in search of a construct. *Perspectives on Psychological Science*, 16(2), 396-414.  
<https://doi.org/10.1177/1745691619863798>
- The Sentencing Project. (2021). *Trends in U.S. corrections*.  
<https://www.jstor.org/stable/resrep32756>
- Shriberg, D., Song, S. Y., Halsell Miranda, A., & Radliff, K. M. (2013). *School psychology and social justice: Conceptual foundations and tools for practice*. Routledge.
- Sisense. (n.d.). What baby names tell us about ethnic and gender trends.  
<https://cdn.sisense.com/wp-content/uploads/What-Baby-Names-Tell-Us-About-Ethnic-and-Gender-Trends.pdf>
- Skiba, R. (2014). The failure of zero tolerance. *Reclaiming Children and Youth*, 22(4), 27-34.
- Skiba, R., Arredondo, M., & Williams, N. (2014). More than a metaphor: The contribution of exclusionary discipline to a school-to-prison pipeline. *Equity & Excellence in Education* 47(4), 546-564. <https://doi.org/10.1080/10665684.2014.958965>

- Skiba, R. J., & Leone, P. E. (2002). *Profiled and punished: How San Diego schools undermine Latino and African American student achievement* (Report No. ED471936). Applied Research Center. <http://files.eric.ed.gov/fulltext/ED471936.pdf>
- Sleek, S. (2018, January 31). *The bias beneath two decades of measuring implicit bias*. Association for Psychological Science. <https://www.psychologicalscience.org/observer/the-bias-beneath-two-decades-of-measuring-implicit-associations>
- Smolkowski, K., Girvan, E. J., McIntosh, K., Nese, R. N. T., & Horner, R. H. (2016). 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. <https://doi.org/10.17988/bedi-41-04-178-195.1>
- Smylie, M., & Murphy, J. (2018). School leader standards from ISLLC to PSEL: Notes on their development and the work ahead. *UCEA Review*, 24-28. <http://npbea.org/wp-content/uploads/2018/10/Smylie-Murphy-2018-School-leader-standards-from-ISLLC-to-PSEL-UCEA-Review-article-only.pdf>
- Social Security Administration. (n.d.). *Get ready for baby*. <https://www.ssa.gov/OACT/babynames/index.html>
- Staats, C. (2014, May). *Implicit racial bias and school discipline disparities: Exploring the connection*. Kirwan Institute. <https://kirwaninstitute.osu.edu/sites/default/files/documents/ki-ib-argument-piece03.pdf>
- Starcke, M., & Porter, S. R. (2019). Do student conduct administrators discriminate against black students? An analysis of drug sanctions using vignettes. *Review of Higher Education*, 42(2), 765-792. <https://doi.org/10.1353/rhe.2019.0014>



- Stoet, G. (n.d.). *Implicit Association Task (IAT) to measure racial discrimination*. PsyToolkit.  
<https://www.psychtoolkit.org/experiment-library/iat2.html>
- Suh, S., & Suh, J. (2007). Risk factors and levels of risk for high school dropouts. *Professional School Counseling, 10*(3), 297-306.  
<https://doi.org/10.5330/prsc.10.3.w26024vww6541gv7>
- Sullivan, G., & Artino, A. R., Jr. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education, 5*(4), 541-542. <https://doi.org/10.4300/jgme-5-4-18>
- Swearingen, J. M. (2014). *Principal longevity, leadership behaviors, and student academic achievement* (Publication No. 3672977). [Doctoral dissertation, University of South Carolina]. ProQuest & Theses Global.
- Taylor, T. O., & Bailey, T. M. (2022). Does race matter? An experimental vignette study on harm severity, college student discipline, and restorative justice. *Journal of Diversity in Higher Education, 17*(3), 385-401. <https://doi.org/10.1037/dhe0000427>
- Tosi, J. D., Ledesma, R. D., Díaz Lázaro, C. M., & Poó, F. M. (2020). Implicit attitudes towards risky driving behaviors: Evidence of validity for the implicit association test. *Journal of Safety Research, 75*, 284-291. <https://doi.org/10.1016/j.jsr.2020.08.008>
- U.S. Department of Education Office for Civil Rights. (2014). *Civil rights data collection: Data snapshot: School discipline*. <https://www2.ed.gov/about/offices/list/ocr/docs/crdc-discipline-snapshot.pdf>

- Van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M., & Holland, R. W. (2010). The implicit prejudiced attitudes of teachers: Relations to teacher expectations and the ethnic achievement gap. *American Educational Research Journal*, 47(2), 497-527.  
<https://doi.org/10.3102/0002831209353594>
- Weisburst, E. K. (2019). Patrolling public schools: The impact of funding for school police on student discipline and Long-term education outcomes. *Journal of Policy Analysis and Management*, 38(2), 338-365. <https://doi.org/10.1002/pam.22116>
- Welch, K., & Payne, A. A. (2018). Zero tolerance school policies. In J. Deakin, E. Taylor, & A. Kupchik (Eds.), *The Palgrave international handbook of school discipline, surveillance, and social control* (1st ed., pp. 215-234). Palgrave Macmillan.  
<https://doi.org/10.1007/978-3-319-71559-9>
- Williams, J., Lewis, C., Starker Glass, T., Butler, B. R., & Hoon Lim, J. (2020). The discipline gatekeeper: Assistant principals' experiences with managing school discipline in urban middle schools. *Urban Education*, 1-29. <https://doi.org/10.1177/0042085920908913>
- Young, B. L., Madsen, J., & Young, M. A. (2010). Implementing diversity plans: Principals' perception of their ability to address diversity in their schools. *NASSP Bulletin*, 94(2), 135-157. <https://doi.org/10.1177/0192636510379901>
- Young, M. (2018). *NELP standards approved by CAEP and NPBEA*. University Council for Educational Administration. [https://ucea.org/p\\_improve\\_nelp\\_standard.php](https://ucea.org/p_improve_nelp_standard.php)
- Young, M., & Perrone, F. (2016). How are standards used, by whom, and to what end? *Journal of Research on Leadership Education* 11(1), 3-11.  
<https://doi.org/10.1177/1942775116647511>

APPENDIX A  
CONTENT PERMISSIONS



Miller, Catherine <Catherine.Miller@avila.edu>

Sun 3/6/2022 11:07 AM

To: Reshel, Bailey

Bailey,  
Sure. Thanks for reaching out.  
Good luck with your dissertation.  
Cathy Miller



---

**From:** Reshel, Bailey <resh5743@bears.unco.edu>

**Sent:** Sunday, March 6, 2022 10:18 AM

**To:** Miller, Catherine <Catherine.Miller@avila.edu>

**Subject:** (EXTERNAL MSG) Permission to Use Article Figure

Hi Dr. Miller,

I'm a doctoral student at the University of Northern Colorado. I'm reaching out to ask if it would be okay to use the graphic below in my dissertation literature review with appropriate citation. I'll be studying implicit bias in principal decision making.

APPENDIX B  
DEMOGRAPHIC SURVEY QUESTIONS

## DEMOGRAPHIC SURVEY QUESTIONS

Please select your gender:

- ☐ Female
- ☐ Male
- ☐ Non-binary / third gender
- ☐ Prefer not to say

Please select your race:

- ☐ Asian
- ☐ Black or African American
- ☐ Hispanic or Latinx
- ☐ Native American or Alaskan Native
- ☐ Native Hawaiian or Pacific Islander
- ☐ White or European American

What is your current job title?

- ☐ Principal
- ☐ Assistant Principal
- ☐ Vice Principal
- ☐ Dean

For how many years have you worked at your current school?

- ☐ 0-3
- ☐ 4-6
- ☐ 7-9
- ☐ 10+

Please estimate the number of hours of diversity training you have engaged in over the past three years (i.e., 2021, 2022, 2023). Trainings may have occurred in higher education classes, as professional development trainings, or as continuing education credits. Diversity trainings include any type of instruction that intended to:

- ☐ Foster cultural responsiveness
- ☐ Provide information on effectively interacting with diverse populations
- ☐ Provide information on implicit and/or explicit biases
- ☐ Encourage self-reflection of personal biases and/or personal cultural identity/intersectionality
- ☐ Reduce prejudice and discrimination
- ☐ Challenge preexisting knowledge about other's experiences and identities
- ☐ Provide information on how to encourage positive group interactions with dissimilar others

- Provide information on dismantling systems of oppression
  - 0-3 hours
  - 4-6 hours
  - 7-9 hours
  - 10-12 hours
  - 13+ hours

Select all the types of disciplinary consequences you've used with students over the past three years:

- ☐ Restorative practice
- ☐ Detention
- ☐ In school suspension
- ☐ Out of school suspension
- ☐ Expulsion
- ☐ Corporal punishment

APPENDIX C  
DISCIPLINE VIGNETTES



## DISCIPLINE VIGNETTES

Survey Group 1	Survey Group 2	Survey Group 3
Brian was talking during independent work time. When I asked him to stop because I should not be able to hear him all the way across the room, he told me to “just ignore it then, everyone else is”.	Darius was talking during independent work time. When I asked him to stop because I should not be able to hear him all the way across the room, he told me to “just ignore it then, everyone else is”.	Jacob was talking during independent work time. When I asked him to stop because I should not be able to hear him all the way across the room, he told me to “just ignore it then, everyone else is”.
This is the second time this week Darius has told me an assignment is stupid and not worth his time. I’ve asked to talk privately about what’s getting in the way and maybe to talk about alternative options for him and he just says the work is stupid in a rude tone.	This is the second time this week Jacob has told me an assignment is stupid and not worth his time. I’ve asked to talk privately about what’s getting in the way and maybe to talk about alternative options for him and he just says the work is stupid in a rude tone.	This is the second time this week Brian has told me an assignment is stupid and not worth his time. I’ve asked to talk privately about what’s getting in the way and maybe to talk about alternative options for him and he just says the work is stupid in a rude tone.
Jacob slammed his book down during independent reading and yelled, “this is effing stupid!”. I reminded him of how much time was left (just 5 minutes) and to be respectful of his classmates until then. He put headphones in and turned his music up loud enough that the whole class could hear. This is my fourth interaction like this with him this semester.	Brian slammed his book down during independent reading and yelled, “this is effing stupid!”. I reminded him of how much time was left (just 5 minutes) and to be respectful of his classmates until then. He put headphones in and turned his music up loud enough that the whole class could hear. This is my fourth interaction like this with him this semester.	Darius slammed his book down during independent reading and yelled, “this is effing stupid!”. I reminded him of how much time was left (just 5 minutes) and to be respectful of his classmates until then. He put headphones in and turned his music up loud enough that the whole class could hear. This is my fourth interaction like this with him this semester.

APPENDIX D  
RACE IMPLICIT ASSOCIATION TASK

## RACE IMPLICIT ASSOCIATION TASK

Block instructions:

<b>WHITE NAME</b>	<b>BLACK NAME</b>
<p>Powered by Qualtrics <a href="#">↗</a></p> <p style="font-size: 2em; font-weight: bold;">+</p> <p>Instructions: Place your left and right index fingers on the E and I keys. At the top of the screen are 2 categories. In the task, words and/or images appear in the middle of the screen.</p> <p>When the word/image belongs to the category on the left, press the <b>E</b> key as fast as you can. When it belongs to the category on the right, press the <b>I</b> key as fast as you can. If you make an error, a red <b>X</b> will appear. Correct errors by hitting the other key.</p> <p>Please try to go as <i>fast as you can</i> while making as few errors as possible.</p> <p>When you are ready, please press the [Space] bar to begin.</p> <p>Part 1 of 7</p>	

<b>unpleasant</b>	<b>pleasant</b>
<p>Powered by Qualtrics <a href="#">↗</a></p> <p style="font-size: 2em; font-weight: bold;">+</p> <p>Now, the categories have changed, but the rules remain the same. Please try to go as <i>fast as you can</i> while making as few errors as possible. Correct errors by hitting the other key.</p> <p>When you are ready, please press the [Space] bar to begin.</p> <p>Part 2 of 7</p>	

<b>WHITE NAME</b> or <b>unpleasant</b>	<b>BLACK NAME</b> or <b>pleasant</b>
<p>Powered by Qualtrics <a href="#">↗</a></p> <p>+</p> <p>Privacy <a href="#">↗</a> &amp; Terms <a href="#">↗</a></p> <p>Now the four categories you saw separately will appear together. Remember, each word/image fits in only one of the four categories.</p> <p>Use the <b>E</b> key for the two categories on the left and the <b>I</b> key for the two categories on the right. Again, try to go as fast as possible without making mistakes. Correct errors by hitting the other key. Practice this combination now.</p> <p>When you are ready, please press the [Space] bar to begin.</p> <p>Part 3 of 7</p>	

<b>BLACK NAME</b>	<b>WHITE NAME</b>
<p>Powered by Qualtrics <a href="#">↗</a></p> <p>+</p> <p>Privacy <a href="#">↗</a> &amp; Terms <a href="#">↗</a></p> <p>Notice the categories from before have switched sides. Please practice this new configuration now. Remember to try to go as <i>fast as you can</i> while making as few errors as possible. Correct errors by hitting the other key.</p> <p>When you are ready, please press the [Space] bar to begin.</p> <p>Part 5 of 7</p>	

<b>BLACK NAME</b> or <b>unpleasant</b>	<b>WHITE NAME</b> or <b>pleasant</b>
<p>Powered by Qualtrics <a href="#">🔗</a></p> <p><b>+</b></p> <p><a href="#">Privacy</a> <a href="#">🔗</a> &amp; <a href="#">Terms</a> <a href="#">🔗</a></p> <p>Notice the four categories have been combined again, but in a new configuration. Please practice this combination now, and remember to go as fast as you can while making as few mistakes as possible. Correct errors by hitting the other key.</p> <p>When you are ready, please press the [Space] bar to begin.</p> <p>Part 6 of 7</p>	

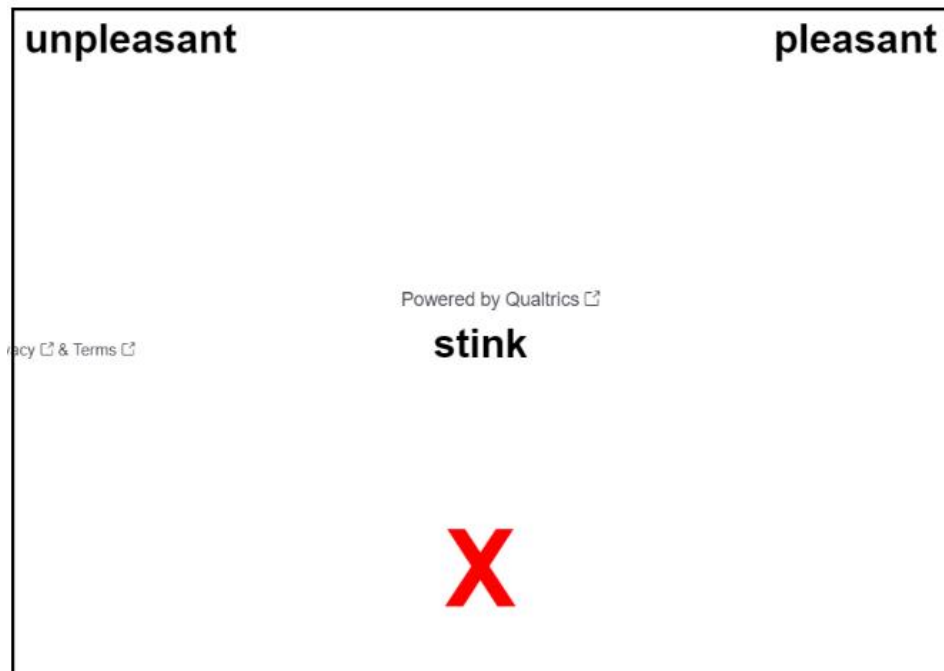
Activity examples:

<b>WHITE NAME</b>	<b>BLACK NAME</b>
<p>Powered by Qualtrics <a href="#">🔗</a></p> <p><b>RASHAUN</b></p> <p><a href="#">Privacy</a> <a href="#">🔗</a> &amp; <a href="#">Terms</a> <a href="#">🔗</a></p>	

WHITE NAME or unpleasant	BLACK NAME or pleasant
<p>Powered by Qualtrics <a href="#">↗</a></p> <p><b>lucky</b></p> <p><a href="#">Privacy</a> <a href="#">↗</a> &amp; <a href="#">Terms</a> <a href="#">↗</a></p>	

Corrective feedback examples:

WHITE NAME	BLACK NAME
<p>Powered by Qualtrics <a href="#">↗</a></p> <p><b>RASHAUN</b></p> <p><b>X</b></p> <p><a href="#">Privacy</a> <a href="#">↗</a> &amp; <a href="#">Terms</a> <a href="#">↗</a></p>	



APPENDIX E

INSTITUTIONAL REVIEW BOARD APPROVAL LETTER





Date: 09/16/2022  
 Principal Investigator: Bailey Reshel  
 Committee Action: **IRB EXEMPT DETERMINATION – New Protocol**  
 Action Date: 09/16/2022  
 Protocol Number: [2205038476](#)  
 Protocol Title: Combating The Racial Discipline Gap: Does Diversity Training Mitigate Implicit Bias in The Discipline Decision-Making Process?  
 Expiration Date:

The University of Northern Colorado Institutional Review Board has reviewed your protocol and determined your project to be exempt under 45 CFR 46.104(d)(702) for research involving

Category 2 (2018): EDUCATIONAL TESTS, SURVEYS, INTERVIEWS, OR OBSERVATIONS OF PUBLIC BEHAVIOR. Research that only includes interactions involving educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior (including visual or auditory recording) if at least one of the following criteria is met: (i) The information obtained is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained, directly or through identifiers linked to the subjects; (ii) Any disclosure of the human subjects' responses outside the research would not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, educational advancement, or reputation; or (iii) The information obtained is recorded by the investigator in such a manner that the identity of the human subjects can readily be ascertained, directly or through identifiers linked to the subjects, and an IRB conducts a limited IRB review to make the determination required by 45 CFR 46.111(a)(7).

You may begin conducting your research as outlined in your protocol. Your study does not require further review from the IRB, unless changes need to be made to your approved protocol.

**As the Principal Investigator (PI), you are still responsible for contacting the UNC IRB office if and when:**



- You wish to deviate from the described protocol and would like to formally submit a modification request. Prior IRB approval must be obtained before any changes can be implemented (except to eliminate an immediate hazard to research participants).
- You make changes to the research personnel working on this study (add or drop research staff on this protocol).
- At the end of the study or before you leave The University of Northern Colorado and are no longer a student or employee, to request your protocol be closed. \*You cannot continue to reference UNC on any documents (including the informed consent form) or conduct the study under the auspices of UNC if you are no longer a student/employee of this university.
- You have received or have been made aware of any complaints, problems, or adverse events that are related or possibly related to participation in the research.

If you have any questions, please contact the Research Compliance Manager, Nicole Morse, at 970-351-1910 or via e-mail at [nicole.morse@unco.edu](mailto:nicole.morse@unco.edu). Additional information concerning the requirements for the protection of human subjects may be found at the Office of Human Research Protection website - <http://hhs.gov/ohrp/> and <https://www.unco.edu/research/research-integrity-and-compliance/institutional-review-board/>.

Sincerely,

A handwritten signature in blue ink that reads "Nicole Morse".

Nicole Morse  
Research Compliance Manager

University of Northern Colorado: FWA00000784