



VCU

Virginia Commonwealth University
VCU Scholars Compass

Theses and Dissertations


Graduate School

2020

IMPACT OF STUDENT'S GENDER AND PERCEIVED SKIN TONE ON EDUCATORS' DISCIPLINARY DECISIONS

Kierstyn K. Johnson-Wigfall
Virginia Commonwealth University

Follow this and additional works at: <https://scholarscompass.vcu.edu/etd>

 Part of the [Accessibility Commons](#), [Educational Leadership Commons](#), and the [Other Education Commons](#)

© The Author

Downloaded from

<https://scholarscompass.vcu.edu/etd/6214>

This Dissertation is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

IMPACT OF STUDENT'S GENDER AND PERCEIVED SKIN TONE ON EDUCATORS'
DISCIPLINARY DECISIONS

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of
Philosophy at Virginia Commonwealth University

by

Kierstyn K. Johnson-Wigfall
B.A., Hampton University, 2007
M.S., Howard University, 2011

Director: Charol Shakeshaft, Ph.D.

Professor, Department of Educational Leadership
School of Education
April 2020

Virginia Commonwealth University
Richmond, Virginia

Acknowledgments

I would like to thank the many people in my life that have helped me make it to this stage in my life. First I would like to thank the three women in my life who have reared me to be the woman I am today. To my mother, who always encouraged me and read all of my writings. You truly have unconditional love for me, and you never give up on me. You have taught me to have patience with others and helped me understand the importance of receiving an education. Thank you for every tuition payment you made! I hope I made you happy! I pray that I will be half the mother that you are.

To my Aunt Bev, also known as my "Favorite Aunt," you taught me perseverance. I watched you work endlessly on home improvement and auto projects. You showed me that hard work pays off, and you gave me my work ethic (Remember when you made me fix the roof with you? We worked for ten hours straight without a break! I thought I was going to die!). You taught me to push myself whenever I felt like giving up.

Last but not least, my grandmother. Although she was called to heaven by God during my first semester of the Ph.D. program, she was instrumental in my success. She taught me lessons that I reflect upon to this day. She always told me that "nothing good comes easy" and that "once acquired, no one can take your education away from you."

To my chair, Dr. Charol Shakeshaft, thank you for helping me make my research vision a reality. You helped me appreciate the research process and made sure that I stayed on track. There are no words I could use to thank you. Additionally, I would like to thank my committee members, Dr. Nora Alder, Dr. Yvonne Brandon, and Dr. Mansfield. Thank you for giving me feedback to make this dissertation complete. Your input was priceless!

Table of Contents

	Page
Acknowledgments.....	ii
List of Tables.....	vii
List of Figures.....	x
Abstract.....	xi
Chapter 1. Introduction.....	1
Historical Foundations.....	3
Manifestation of Colorism	4
Purpose of the Study	6
Theoretical Framework.....	7
Problem Statement	8
Research Questions	9
Overview of the Methodology.....	10
Chapter 2. Review of the Literature.....	11
Literature Review Methodology	11
Disproportionality	12
Educator's Perception of African American Students.....	18
Skin Tone Bias.....	35
Skin Tone and Job Selection.....	44
Skin Tone and Adoption.....	45
Skin Tone and Marriage.....	46
Skin Tone and Classroom Experience.....	46

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Skin Tone and Educational Achievement.....	48
Skin Tone and Teacher Preparation	49
Skin Tone and School Suspension.....	51
Skin Tone and Transitioning.....	53
Skin Tone and the United States Penal System.....	54
Skin Tone and Perceived Intelligence	57
Own Race Bias.....	58
Summary.....	60
Purpose.....	61
Chapter 3. Methodology.....	63
Methodological Approach.....	64
Population and Sampling.....	64
Variables.....	65
Treatment Variables.....	65
Validity of Treatment Variables.....	68
Dependent Variable.....	71
Independent Variable.....	73
Post Pilot Study to the Survey	74
Data Collection.....	75
Data Analysis.....	76
School Discipline Scenarios.....	78
Practice Question	78
Defiance of Authority/Insubordination.....	78

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Classroom and Campus Disruption	78
Disruptive Demonstrations.....	79
VCU IRB and School Division Research Review Board	80
Limitations.....	80
Delimitations.....	81
Assumptions.....	82
Researcher's Perspective.....	82
Disclosures.....	83
Chapter 4. Findings.....	84
Research Question 1.....	86
Research Question 2.....	87
A. Comparison of All Skin Tones and Control Group.....	87
B. Comparison of All Skin Tones Without Control Group.....	87
C. Comparison of Light, Medium and Dark Students With the Control Group.....	87
D. Comparison of Light, Medium and Dark Students Without the Control Group.....	88
E. Comparison of White Students with the Control group.....	88
F. Comparison of Students of Color, White Students and Control Group.....	89
G. Comparison of Students of Color and Control Group.....	90
H. Comparison of Students of Color and White Students.....	90
I. Comparison of All Students and Control Group.....	91
Research Question 3.....	92
Research Question 4.....	93
A. Comparison of Discipline Decisions for Students of Color, White Students and Control Group by Intersection of Gender.....	93

B. Comparison of Discipline Decisions for Students of Color by Gender.....	93
Research Question 5.....	95
Research Question 6.....	95
Research Question 7.....	96
Research Question 8.....	98
Research Question 9.....	100
Research Question 10.....	101
Research Question 11.....	104
Chapter 5. Discussion, Conclusion, and Recommendations.....	107
Discussion of Findings.....	107
Racial Mechanics Model: CRT and Implicit Social Cognition.....	110
Current State of Affairs- Policy and Political	114
Conclusion.....	116
Recommendations.....	117
References.....	120
Appendix A.....	133
Appendix B.....	140
Appendix C.....	141
Appendix D.....	142
Appendix E	143
Appendix F.....	144
Appendix G.....	145

List of Tables

Table	Page
3.2 Research Questions Matrix	77
4.1 Participant Gender.....	85
4.2 Participants Race.....	85
4.3 Frequency of Discipline.....	86
4.4 Sample Size, Mean and Standard Deviation for Student Skin Tone and Discipline Decision.....	89
4.5 Sample Size, Mean and Standard Deviation for Student Skin Tone and Control Group...	89
4.6 Sample Size, Mean and Standard Deviation for Student of Color and Control Group.....	90
4.7 Sample Size, and Standard Deviation for Student of Color and White Students.....	91
4.8 Sample size, Mean and Standard Deviation for All Students and Control Group.....	91
4.9 Summary of Statistically Significant Differences for Question Two-Part A through H....	92
4.10 Sample Size, Mean and Standard Deviation for Student Gender and Discipline Decision	92
4.11 Summary of Statistically Significant Difference for Question Three.....	93
4.12 Sample Size, Mean and Standard Deviation for Student Skin Tone/Gender and “Harshness” of Discipline Decision.....	94
4.13 Summary of Statistically Significant Differences for Questions Four Part A and Part B.	95
4.14 Sample, Mean and Standard Deviation by Race of Educator and Discipline Decision...	96
4.15 Sample, Mean and Standard Deviation by Race of Educator and Student Skin Tone for Scenario 1.....	96
4.16 Sample, Mean and Standard Deviation by Race of Educator and Student Skin Tone for Scenario 2.....	97

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

4.17 Sample, Mean and Standard Deviation by Race of Educator and Student Skin Tone for Scenario 3.....	97
4.18 Sample, Mean and Standard Deviation by Race of Educator and Student Skin Tone for Scenario 4.....	97
4.19 Summary of Statistically Differences for Question Seven.....	98
4.20 Sample, Mean and Standard Deviation by Race of Educator and Student Gender for Scenario 1.....	98
4.21 Sample, Mean and Standard Deviation by Race of Educator and Student Gender for Scenario 2.....	99
4.22 Sample, Mean and Standard Deviation by Race of Educator and Student Gender for Scenario 3.....	99
4.23 Sample, Mean and Standard Deviation by Race of Educator and Student Gender for Scenario 4.....	100
4.24 Summary of Statistically Differences for Question Eight.....	100
4.25 Sample, Mean and Standard Deviation for Educator Gender and Discipline Decision	101
4.26 Summary of Statistically Significant Differences for Question Nine.....	101
4.27 Sample, Mean and Standard Deviation by Gender of Educator and Student Skin Tone for Scenario 1.....	102
4.28 Sample, Mean and Standard Deviation by Gender of Educator and Student Skin Tone for Scenario 3.....	102
4.29 Sample, Mean and Standard Deviation by Gender of Educator and Student Skin Tone for Scenario 4.....	102
4.30 Sample, Mean and Standard Deviation by Gender of Educator and Student Skin Tone for Scenario 2.....	102
4.31 Summary of Statistically Significant Differences for Question Ten.....	103
4.32 Sample, Mean and Standard Deviation by Gender of Educator and Student Gender for Scenario 1.....	104
4.33 Sample, Mean and Standard Deviation by Gender of Educator and Student Gender for Scenario 2.....	104

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

4.34 Sample, Mean and Standard Deviation by Gender of Educator and Student Gender for Scenario 3.....	105
4.35 Sample, Mean and Standard Deviation by Gender of Educator and Student Gender for Scenario 4.....	105
4.36 Summary of Statistically Significant Differences for Question Ten.....	105
4.37 Summary of Statistically Significant and Meaningful Results with Interpretation.....	106

List of Figures

Figure	Page
1.1 Racial Mechanics Model.....	8
3.1 Original Student Picture Matrix.....	68
3.2 Final Student Picture Matrix.....	70
3.3 Massey-Martin Scale.....	71

Abstract

IMPACT OF STUDENTS' SKIN TONE AND GENDER ON EDUCATORS' DISCIPLINARY DECISIONS

By Kierstyn Johnson-Wigfall

A dissertation prospectus submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University

Director: Charol Shakeshaft, Ph.D., Professor, Department of Educational Leadership

The purpose of this proposed dissertation is to examine the relationship of students' skin color and gender to school disciplinary decisions for African American¹ children. While skin tone bias or colorism² is frequently studied, limited research has been completed about its prevalence in American public schools during the discipline process. Understanding if colorism is a factor that contributes to school disciplinary³ sanctioning is vital because racial disparities influence life trajectories (Skiba, Horner, Rausch, May & Tobin, 2011). For example, school suspension leads to higher rates of absenteeism, lower academic achievement and higher chances of incarceration. Even students who receive an office referral and return to class minutes later miss pertinent instruction time. This study aims to: 1) add to the body of literature about this

¹ In the context of this dissertation the term African American is synonymous for black. This decision goes along with the wording used in the United States Census.

² Skin tone bias and colorism are used interchangeably in the context of this dissertation. "Colorism is the allocation of privilege and disadvantage according to the lightness or darkness of one's skin" (Burke, p. 17, 2008), and tends to favor lighter skin over darker skin due to an appearance closer to a white phenotype (Hall, 2005). Hair texture, eye color, and facial features, as well as education and income can impact perceptions of skin tone (Hunter, 2005).

³ In the context of this dissertation the term discipline refers to office referral and suspensions unless other interpretations apply.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

topic, 2) shape political discourse about future school discipline policies and procedures, and 3) assist school divisions in designing professional development to bring awareness to the idea of colorism in the field of education.

The literature shows that African American students are disciplined at a higher rate than children of other races⁴. Moreover, dark-skinned African American children are disciplined at a higher rate than light-skinned African American children. African American male and female students are more likely to receive an office referral and subsequently suspended than other race peers. A student's skin tone and gender could impact office referrals and suspension rates.

Using implicit social cognition⁵ and Critical Race Theory (CRT)⁶ as the theoretical framework, this study will seek to understand the school personnel's disciplinary decisions based on the skin tone and gender of students, that may occur outside of conscious awareness or control. This study will add to the current body of literature by gaining a better understanding of factors that can predict how school personnel discipline African American students of varying skin tones, compared to white students.

⁴ The scientific consensus is that race does not exist as a biological category among humans. However, in American society, race is distinguished by physical characteristics. (Clair & Denis, 2015)

⁵ Rooted in implicit learning theory that describes how much of our knowledge can be acquired and shaped outside of awareness and without intention (Chao & Willab, 2007, p. 681). Cleeremans (2003, p. 491) defines implicit learning as "the process through which we become sensitive to certain regularities in the environment (1) in the absence of intention to learn about those regularities, (2) in the absence of awareness that one is learning, and (3) in such a way that the resulting knowledge is difficult to express".

⁶ Critical Race Theory (CRT) developed in the mid-1970s, as a number of lawyers, activists, and legal scholars across realized the advances of the civil rights era of the 1960s had stalled or reversed. Researchers realized new theories and strategies were needed to combat the subtler forms of racism that were gaining ground (Delgado & Stefanie, 2017).

Chapter 1: Introduction

The current state of educational affairs in the United States consists of many issues that need to be resolved. For example, for decades, the major problem in education has been closing the achievement gap between white and black students. Researchers have noted that the school discipline gap needs to close before the achievement gap can be diminished, and yet the focus on the achievement gap seldom includes an analysis of disparities in discipline. Gregory, Skiba & Noguera (2010) synthesized the research on academic achievement and discipline, and concluded that it is not difficult to conceptualize the impact frequent suspensions have on black students, or any population of students, who are frequently suspended or expelled. The disciplinary practices that most schools in the United States employ perpetuate the achievement gap by decreasing the amount of time students spend in class engaging in the curriculum. The issue of the disproportionate suspension of students of color is not new. The Children's Defense Fund introduced the discipline gap in 1975 by publishing a report, which chronicled that black students were two to three times overrepresented in school suspensions compared to white students in proportion to enrollment rates in localities across the nation. Over forty years later, the data proves the same disparity exists.

Research in this area has slowly progressed. Research finds that a discipline gap exists, but fails to investigate the contributing factors in classrooms across America where the school discipline gap not only exists between black and white students but also exists within the black student population. Within the black student population, light-skinned and dark-skinned black students are suspended at a disproportionate rate, (Hannon, Defina & Bruch, 2013), which may

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

indicate a skin tone bias that has long existed in many facets of American society. To better understand skin tone bias, exploring the historical context is useful.

Historical Foundations

That may surprise you, but whites do the same thing.
You have the blond hair and blue eyes and from there
you go all the way down to the Greeks and the Italians.
And they tell you that. See, that is something that
American culture has internalized. It began as black
and white but it's much, much larger than that now
because it's a cultural value. When people discriminate
on the basis of skin color they're simply being American.

~ Ronald E. Hall, Star Tribune, Feb 6, 1993

The United States of America is a country considered by many to be the land of opportunity where ambition, determination, and hard work may result in any man or woman achieving his or her dreams. The classic essay, *A Nation of Immigrants* by President John F. Kennedy, details the journey of the Irish, German, and Scandinavian people to the New World. Included in the most recent edition of the work is an introduction by Senator Edward M. Kennedy, who reflected on his brother's concern for immigration reform. He recalled President Kennedy stating, "This was the secret of America: A nation of people with the fresh memory of old traditions who dared to explore new frontiers, people eager to build lives for themselves in a spacious society that did not restrict their freedom of choice and action" (Kennedy, 1964, p. 9). This statement, while extremely idealistic and very patriotic on the surface, falls short in acknowledging the history and experience of all American citizens. It primarily captures the perspectives of people of European descent and universalizes them, assuming that all ethnic groups in the United States are comprised of voluntary immigrants who share similar circumstances. Native Americans, the only indigenous inhabitants of the land now known as the United States, were here before European settlers, yet were displaced from their homes and territories, relocated to reservations as a result of the doctrine of Manifest Destiny. By contrast, the peoples of West Africa were taken captive against their will and brought to the U.S. to

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

manifest the destiny of the southern American economy as chattel slaves. Although every race and ethnic group can trace their earliest ancestral origins to the continent of Africa, (Gibbons, 2006). President Kennedy's chronicle passes over the history of Native and African Americans under the premise that every U.S. citizen can trace roots back to voluntary immigrant status and that all Americans had equal access to freedom of choice. People of color have been at the forefront of oppression, discrimination and marginalization since the genesis of their interaction with settlers of European descent.

Although all people of color have faced and still face injustices in the United States, the inequalities African Americans have encountered as a result of the system of slavery have been uniquely disturbing and impactful. One of the facets of this institution that was so impairing was the categorical denial of educational opportunities. Enslavers prohibited literacy for the slaves on their plantations as it was believed that an educated slave would not remain a slave for long. Hence, in many Confederate states, it was illegal for a person of African descent to learn to read. Beginning in the year 1619, human beings from the continent of Africa were taken from their native land by way of the Atlantic Slave Trade and held in bondage for roughly 246 years in America. Chattel slavery stripped them of their culture, language, traditions, and freedom as they were forced to leave the only home they ever knew. Initially, during the earliest periods of American colonialism, some enslavers granted manumission to their African thralls if they converted to Christianity. However, after 1667, the enslaved or free status of Africans became linked to skin color.

The Manifestation of Colorism

To justify the inhumane exploitation as a societal necessity and to support legal enslavement, champions of the slaveholding culture employed propaganda to initiate a black

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

inferiority campaign to portray Africans in an unfavorable light. Africans were depicted negatively, and skin color was the sole factor. “To justify permanent enslavement of Africans and the horrific treatment they received, the colonist began to publicize grotesque depictions of Africans to foster fear and loathing and to associate dark-skinned people with the cursed descendants of Ham⁷” (Norwood, p. 14, 2014). The slave trade was supported because European Americans thought that the institution was a contemporary manifestation of the Curse of Ham (in conjunction with being profitable). Pseudo-science was used to rationalize the inferior treatment of persons of African descent.

Colorism can be seen in different countries. In America, colorism started as a result of the rape of African female slaves by their white male captors. Mixed-raced slaves were born out of these relations, and a new class developed, and slaves were divided by skin tone (previously divided by language). Most slave owners gave slaves with lighter complexions preferential treatment by giving them jobs inside of the plantation owners' homes instead of strenuous jobs outside in the plantation fields. The offspring were called “mulattoes,” and they were lighter than those born to two African parents, and sometimes parallel in complexion to white people. Over time, interracial stratification occurred, as lighter-skinned blacks were believed by whites to be more visually attractive and intellectually superior to their darker-skinned compatriots (Norwood, 2014)

Colorism can be observed in many aspects of American society (Coard, Breland & Raskin, 2001). For example, many light-skin blacks hold higher-level faculty positions at the university compared to dark-skin blacks (Norwood, 2014). Research has shown that dark-

⁷ The story of Ham is mentioned in the bible (Genesis 10:6). Descendants of Ham were cursed because he saw his father's naked body and failed to cover him. It is implied in the bible that descendants of Ham had dark-skin and lived in regions in Africa.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

skinned blacks that are arrested receive longer prison sentences for similar crimes (Eberhardt, Davies, Purdie-Vaughns & Johnson, 2006) than light-skinned blacks. Colorism also impacts the likelihood of marriage for women (Hamilton, Goldsmith & Rarity, 2009). White, bi-racial, and light-skinned black children are more likely to be adopted than dark-skinned black children (Fedders, 2010).

Purpose of the study

According to the report, *Black Lives Matter: The Schott 50 State Report on Public Education and Black Males* (2015), the 2012-2013 national graduation rate for black males was 59 percent, compared to the higher rates of 65 percent for Latino males, and 80 percent for white males. Discipline issues have been shown to have a negative impact on the graduation rates for all students. The report found that 15 percent of black male students received out-of-school suspensions, while the same was true for 7 percent of their Latino peers and five percent of white counterparts. African American male youth are three times more likely to be victims of reported child abuse or neglect, three times more likely to be victims of robbery, and five times more likely to be victims of homicide. Of black children within the range of grades 7-12, 35 percent have been suspended or expelled at some point during their tenure as students compared to 20 percent of Hispanics and 15 percent of whites.

The educational and disciplinary outcomes are better for African American girls than boys. However, African American girls are suspended at higher rates (12 percent) than girls of any other race or ethnicity. Additionally, African American girls are graduating at lower rates than white boys and girls (Office of Civil Rights, 2014). More research is needed to understand if a skin tone bias is impacting school discipline rates for both male and female African American students.

Theoretical Framework

This proposed dissertation will analyze student skin tone and gender to understand the prevalence of skin tone bias or colorism in the classroom, utilizing implicit social cognition and CRT as the theoretical framework. Implicit social cognition is a construct that describes unconscious beliefs. Implicit bias occurs involuntarily without a person's awareness. All human beings have unconscious biases of some sort, even those committed to being impartial and fair (Rachlinski, Johnson, Wistrich, & Guthrie, 2009).

An example of this [implicit bias] could be a school administrator who believes he or she is meting out equal punishments for equivalent infractions, when in fact certain student populations are receiving harsher discipline due to the subtle yet powerful influence of the administrators' implicit biases (Staats, 2014, p. 7).

Greenwald and Banaji (1995) used the term implicit social cognition to describe thoughts and feelings that occur outside of conscious awareness or conscious control concerning social psychological constructs— attitudes, stereotypes, and self-concepts.

CRT is an additional tool that can be used to understand the findings. I see the two as intertwined. Implicit social cognition occurs involuntarily without a person's awareness in relation to social psychological constructs-attitudes, stereotypes, and self-concepts (Greenwald & Banaji, 1995). CRT analyzes race and racism as a social disparity that occurs between dominant and marginalized groups (Hiraldo, 2010; DeCuir & Dixson, 2004; Ladson-Billing & Tate, 1995). Legal researcher Kang (2005) proposed the Racial Mechanics Model to explain race as a social construct using CRT and implicit social cognition as a theoretical framework (see Figure 1.1).



Figure 1.1 Racial Mechanics Model

This model focuses on the interaction between a “perceiver” and “target.” According to Kang (2005), upon encountering a target individual, the perceiver classifies that individual into a " racial category according to relevant mapping rules provided to us by culture and any specific rules relevant to the context. Once that mapping is performed— typically instantaneously—a set of racial meanings are activated that alters the way that the perceiver interacts with the target."

(p. 45)

Problem Statement

The field of education has evolved since the decision of *Brown v. Board of Education* in 1954. Students are educated in the same classrooms as peers of different ethnic groups. The desegregation of American schools ensured that black children would be afforded equal educational opportunities as white children. A little over sixty years later, during a period in time where the United States has had the first African American president, it seems unimaginable that young black children continue to struggle academically and face disproportionately high disciplinary rates.

African American children have the highest suspension and lowest graduation rates when compared to all other student subgroups. This population of students often receives harsher penalties for similar infractions than do other populations. This research is necessary because excessive school discipline influences students for their entire lives. This study will add to the literature because previous studies used decade-old data sets and interviewer ratings of skin color to analyze the relationship between skin tone and school discipline. Additionally, school personnel's attitudes will be measured implicitly instead of explicitly to decrease the chance of response bias or providing a particular response to appear socially desirable. Implicit measures are less susceptible to response bias because they measure automatic responses elicited by stimuli that are congruent or incongruent with one's attitudes (Siegel, Dougherty, Huber, 2012). Because of America's less than favorable past, racism, colorism and sexism are embedded in the culture and are deemed as undesirable traits. This study aims to understand if the "isms" that are present in society are used by educators when determining student discipline infractions.

Research Questions

The proposed dissertation will use a quantitative methodology to answer the following questions:

1. How "harshly" do educators discipline based on scenario type?
2. Are there differences in educator discipline decisions by color of vignette student skin tone?
3. Are there differences in educator discipline decisions by gender of vignette student?
4. Are there differences in educator discipline decisions by intersection of gender and skintone of vignette student?
5. Are there differences in educator discipline decisions by race of educator?
6. Are there differences in educator discipline decisions by race of educator and skin tone of vignette student?

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

7. Are there differences in the discipline decisions made by black and white educators based upon brown skin tones of light, medium, and dark?
8. Are there differences in educator discipline decisions by race of educator and gender of vignette student?
9. Are there difference in educator discipline decision by gender of educator?
10. Are there differences in educator discipline decisions by gender of the educator and vignette student skin tone?
11. Are there differences in educator discipline decisions by gender of the educator and gender of vignette student?

Overview of Methodology

The research questions were answered by distributing an online survey to 1,016 participants. Participants were graduates of the educational leadership and teaching master's degree, post-master's degree, and doctoral programs at a University in the southeast United States. Participants graduated from the University from 2007 through 2018 and have experience as educators. Each participant was asked to identify themselves as either teacher, principal, assistant principal, support staff or other administrator to participate in the study. Participants read four scenarios about a student committing an infraction and asked to rate how likely they were to select a specific type of punishment for the student. There are a total of nine surveys, one for each student gender (male and female) and skin tone (white, light, medium, and dark) combination. One survey contained discipline scenarios without pictures. Participants were randomly assigned to complete one of the nine surveys.

Chapter 2: Review of the Literature

Literature Review Methodology

The Virginia Commonwealth University (VCU) library was used to access a variety of journals, with inclusion and exclusion criteria selected. To ensure that the research was relevant, the determination was made to ensure that the majority of references used in the literature review would be no more than five years old. However, a few groundbreaking articles from decades ago were necessary in certain circumstances. Scholars frequently use these seminal research articles as the foundation for newer studies. Recent articles would provide the most current information about research being conducted about the areas of interest. The search was refined to include specific journal titles to weed out irrelevant material. Journal articles were examined to determine that they were peer-reviewed. Sources that were frequently cited in multiple articles by multiple authors were selected. Reference lists of articles that were well written, with a clearly articulated methods section, were gathered to help further search of articles about the topic. RefWorks was used to organize the articles that were read. Using RefWorks ensured the inclusion of sources in the reference list used to complete this dissertation.

Databases searched: Ebscohost, PsycInfo, Education Research Complete, ERIC, Science Direct, Proquest

Keyword search: cognitive consistency, color bias, colorism, colorism in education, criminal justice, ethnic bias, facial feature bias, implicit association test, implicit social cognition, racial identifiably, racial disparity, racism, racial categorization, school to prison, school discipline disproportionality, school disciplinary procedures and skin tone bias.

Disproportionality (The Black-White Discipline Gap)

In the United States education system, there are large racial disparities in disciplinary practices. The Office for Civil Rights surveyed over 70,000 schools and found that black students are three times more likely to be suspended or expelled than white peers (2014). This phenomenon is referred to as the discipline gap. The issue of the disproportionate number of students of color who are suspended and expelled from American public schools has received national attention. The United States Department of Justice and the United States Department of Education (2014) issued a statement of guidance to assist public schools in following federal law, which prohibits discrimination based on race, color and national origin. The report noted that African American students without disabilities are more than three times as likely as their white peers without disabilities to be expelled or suspended. African-American children comprised "15 percent of students in the Civil Rights Data Collection (CRDC), however they made up 35 percent of students suspended once, 44 percent of those suspended more than once, and 36 percent of students expelled. Additionally, 50 percent of students who were arrested at school or referred to law enforcement are Hispanic or African American" (2014, p. 3). Moreover, racial disparities present in the school discipline data cannot be explained by more serious or frequent behavior by black students.

Using data from the North Carolina Department of Education, Kinsler (2011) asked the question: "How much of the documented discipline gap can be explained by cross-school variation in punishment?" Kinsler (2011) addresses this matter by examining individual punishment decisions within and across schools. The review considered the two discipline margins considered by principals once a referral is received, whether to suspend or not, and the length of the suspension. Participants were sixth ($N=148,444$) and ninth ($N=149,558$) graders.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

This group of students was selected because they would be new to their respective schools and would have never interacted with the administration. The study found that there was a significant statewide racial gap in discipline because of cross-school variation in student punishment. In the Tar Heel state, black 6th grade students were 79 percent more likely to be suspended for violating school rules than white 6th-grade students who violated school rules. Black 9th grade students received 22 percent longer suspensions than white students when committing simple rule violations. Race was found to be related to the initial referral, with more black students being referred for subjective violations (rule violation, aggressive behavior, and rowdiness). The research revealed that the students that were most likely to be suspended were: likely to be male, black, and come from low-income households. Specifically, black students received significantly harsher penalties for the same infractions as white students. The study also investigated how student-teacher and student-principal race interactions affect the likelihood of being referred and the severity of discipline. It found that both within and across schools, there is an increased likelihood of being referred at least once for black students relative to white students regardless of the teacher's race. In North Carolina, the referral gap between black and white students is larger for white teachers than for black teachers, suggesting that instructors differentiated by their respective ethnic backgrounds treat students accordingly. The referral gap between black and white students equals 0.055% points for white teachers and 0.043% points for black teachers. The F-statistic is 1.73, with a p-value of 0.19. Results are similar for student-principal race interactions. The author concluded that although racial bias was not present in the study, it does not mean that it is not present. It may suggest that the race of the teacher or principals plays a minor role. The article was unable to answer why the discipline gap exists within this province.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Skiba, Horner, Chung, Rausch, May and Tobin (2011) analyzed the pattern of office referrals in 364 elementary and middle schools during the 2005-2006 academic year. Discipline data was reported by school personnel through uploading information to a web-based School-wide Information System, which was being used in over 4000 institutions. Using descriptive and logistic regression analyses, the researchers found that African American students are 2.19 (at the elementary level) to 3.78 (at the middle school level) times as likely to receive a discipline referral to the office than their white peers. This research found that the disproportionate representation of African American students in school discipline might occur at either the point of referral or administrative decision. The findings in this study are consistent with the literature and demonstrate the need for changes in policy, practice and research. Unlike other studies, Skiba et al. (2011) research follows students through the disciplinary pipeline (origination of referral and then administrative decision). The results indicated that administrative decision based on the referral significantly contributes to racial disparities in school discipline. The limitation of this study is that it did not describe the cause of the racial discrepancy. More research is needed to understand the casual mechanism for this discrepancy, and then to make recommendations for educators to reduce school discipline rates.

In a theoretical article, Irby (2014) offered a different perspective for the cause of the black-white discipline gap. Using white racial purity desires as the theoretical framework, Irby argued that schools perpetuate a white-supremacist patriarchy to continue the notion of whiteness by continuously eliminating black male bodies using school discipline. That is, removing black male bodies from the school protects the bodily safety and academic achievement of white girls. The author explains that disciplining black males harshly within desegregated schools is a post-civil rights policy based mechanism for maintaining white racial

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

purity in schools. Black males are pushed out of predominately white schools for fear of stigma transference. These schools risk being deemed unbecoming to white culture reproduction.

Schools that are integrated successfully do not perpetuate the cycle of white supremacy and, therefore, are determined to be undesirable and at risk for white-flight abandonment. This study does not explain why black girls are also experiencing higher suspension rates than white girls.

Welch and Payne's (2010) research was scaffolded upon Michael Foucault's notion of the school as an institution aimed at socially controlling students (Foucault, 1977). American educational institutions are increasingly employing a punitive approach for controlling their constituents, with strong parallels to the criminal justice system. For example, Welch and Payne sampled 294 public schools in the United States to test the racial threat⁸ hypothesis to determine if the racial composition of students predicts greater use of punitive controls, regardless of the levels of misbehaviors and delinquency. This investigation analyzed data from the National Study of Delinquency Prevention in Schools. Five degrees of disciplinary response were operationalized using twenty-two questions from a principal questionnaire that asks about hypothetical responses to student misconduct that may be used by school administration. The five degrees were: extreme punitive disciplinary response, zero tolerance⁹, punitive disciplinary response, mild disciplinary response and restitutive. A multivariate analysis produced results that supported the racial threat perspective that schools with a larger percentage of black students are

⁸ Racial threat is a theory that a large or growing minority populations may elicit racialized fears and perceptions of crime (Blalock, 1967). Schools with larger minority student populations may be viewed as more threatening; to address this threat, students in these schools may be punished more severely through suspension and expulsion (Payne and Welch 2010)

⁹ Zero tolerance or zero tolerance policies were created for use in the criminal justice system. Zero tolerance policies were first introduced in American public schools with the passage of the Gun-Free School Act of 1994. These policies initially required a 1-year mandatory expulsion for any student caught with a weapon at school. Later the policies were modified to include a wide variety of offenses that mandated specified consequences for various behaviors

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

more likely to practice extreme punitive discipline and to implement zero tolerance policies. The percentage of African American pupils is significantly and positively related to punitive disciplinary response ($b = .007, p < .001, r^2 = .205$), meaning that schools with a greater percentage of black students are more likely to exact harsh forms of punishment, such as in-school suspensions, detentions, and suspensions. The percentage of black students is significantly and positively related to both the implementation of zero tolerance ($b = .003, p < .001, r^2 = .116$) and the use of extreme punitive disciplinary response ($b = .003, p < .05, r^2 = .132$), meaning that schools with a high percentage of black students are more likely to respond to misbehavior by suspending students for various violations or utilizing zero tolerance policies. Schools with a larger percentage of black students are less likely to use milder forms of punishment, such as counselor visits or oral reprimands. Finally, the percentage of black students is significantly and negatively related to restitutive disciplinary response ($b = -.004, p < .05, r^2 = .186$). Results show that schools with greater percentages of black students are less likely to facilitate practices such as restitution and community service in response to misbehavior. Although this work adds to the literature, the data used in this study was from 1997 and 1998.

Studies about the discipline gap span over a decade. However, the problem still remains that students of color are punished more severely and frequently than their white peers. Milner's (2013) article entitled *Why are Students of Color (Still) Punished More Severely and Frequently Than White Students?* synthesized the body of literature related to this topic and concluded the following: 1) Teachers and administrators may be afraid of their African American and Latina/American students and are likely to be less willing to work with them in order to keep them in the classroom and teach them; 2) Institutional policies on the district and school level (such as zero tolerance) can be rife with racism especially when they are constructed by and follow a

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

white norm that excludes the behavioral and interactional styles and practices of non-white people; 3) Teacher education programs often fail to address the needs of teachers in meeting the needs of their P-12 students by not having a classroom management course; 4) There is disconnect between teachers and students, which can shape what teachers teach, why they teach what they teach, and how they react to their students and what students learn; 5) Leaders do not always provide learning opportunities for teachers to develop, mature, and grow from the disciplinary conflict; 6) There are inadequate counseling and psychological services, that may help educators understand the root of students' discipline problems and provide them with psychological support.

Gregory, Skiba and Noguera (2010) acknowledged that both the discipline gap and the achievement gap have been discussed for over a decade. However, the two topics are often discussed separately. In a theoretical piece, Gregory et al., identified five tasks that should be completed to decrease the discipline and achievement gap simultaneously: Increase the awareness of teachers and administrators of the potential for bias when issuing referrals for discipline, utilize a wide range of consequences in response to behavior problems, treat exclusion of the student as a worse-case scenario option, make an effort to understand the source or the root of the problem, and review the mission of the school with the student when a discipline events. Few studies have examined where and why disproportionality between black and white students is on the increase, especially for black females

In conclusion, the literature uses theoretical frameworks that apply to the criminal justice system. Teachers may not explicitly connect their disciplinary reactions to negative perceptions of black children. However, the immense disproportionality suggests that teachers may be implicitly influenced by stereotypical perceptions ingrained by society that black children need

to be more controlled than white children. Furthermore, the research does not discuss the intraracial disparity within the discipline gap.

Educator Perception of African American Students

In an article published by the New York Times *Where are the Teachers of Color?* (Rich, 2015), noted that “minority” students (comprised of various ethnicities) have collectively become the majority in American public schools. However, the teacher workforce does not mirror the current student demographic. More than 80% of teachers are white. This trend is seen in many cities across the United States. In Boston, there is just one Hispanic teacher for every fifty-two Latina/o students, and one black teacher for every twenty-two African-American students. The ratio of white teachers to white students is one to fewer than three. No one would argue that a teacher of a different race could not teach a child who is of different race. However, as Andres Antonio Alonso, a professor at the Harvard Graduate School of Education noted, “But if there are no people who somehow mirror the parents and the kids, then I think there could be a problem” (Rich, 2015, para. 7). Research has shown that white and black educators view African American students differently.

Francis (2012) used national data from the Early Childhood Longitudinal Study (1998-1999). Participants were kindergarten through eighth grade female students along with their respective English, math and science teachers. Data was used to examine teacher perceptions of student behavior, such as attentiveness and disruptiveness. Teachers were asked the questions: “How often is this student attentive in your class?” and “How often is this student disruptive in your class?” They were given a Likert scale to respond to questions. Teachers were also asked if they would recommend a student for honor classes. The study found that black female students were perceived as less attentive and more disruptive than white, Hispanic and Asian female

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

peers. Additionally, white female students were 19% and Asian female students were 30% more likely to be recommended for honors than black female students. The author noted that the data might reveal evidence of cultural mismatch, which could mean that teachers are misperceiving black girls' behavior as disruptive. Francis concluded that teacher perception is important because it can impact the academic outcomes for black girls. However, further research is needed to understand the mechanisms influencing these perceptions. It is unclear if the results are due to teacher bias or actual student behavior. The article did not discuss the race or gender of the teachers.

Bates and Glick also used data from the Early Childhood Longitudinal Study-Kindergarten Cohort (1998-1999, $n= 21$) to understand the extent to which children receive different evaluations from their teachers depending on the racial/ethnic match of teachers and students (2013). Assessing an individual child by data submitted by multiple teachers, the researchers used PROC MI and PROC MIANALYZE to impute the data. Teachers rated student behavior by using the teacher Social Rating Scale (SRS). SRS asks raters to report the frequency the student: argues, fights, gets angry, acts impulsively, and disturbs on going activities. There were four ratings per child from four different teachers in four different school years. The study found that black students are more likely to be rated as exhibiting problematic behaviors in the classroom. Students also did not receive the same ratings from every teacher. For example, Black children received more favorable ratings from black teachers than white teachers. The authors concluded that black teachers are able to sympathize with black students and look past stereotypes when assessing student behavior. This article did not discuss how teachers perceived students with various skin tones.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Okonofua and Eberhardt (2015) completed two controlled online experiments aimed at understanding how student race may influence teachers' responses to classroom misbehavior. In both studies the researchers showed practicing teachers the records of a black or a white student who misbehaved twice and measured teachers' responses from one infraction to the other. Teachers were asked to report how troubled they would feel as a consequence of the student's misbehavior and how severely they would want to discipline the student. In the second study, teachers were asked to report on the likelihood that they would suspend the student in the future. Findings were compared with the national statistics on school suspension in the United States. The research recruited fifty-seven female kindergarten-through – 12th grade (K-12) teachers (average experience: 14 years; average class size: 26; average age 42 years; 38 white, 2 black, 1 Asian, and 16 unknown). The study had a 2 (student race: black vs. white) x 2 (number of infractions 1 vs. 2) mixed design, with the number of infractions serving as the within subjects factor. In this survey, teachers were shown a picture of middle school and told to imagine themselves as a teacher there. They then reviewed a school record for a student who misbehaved twice. Student race was manipulated by using stereotypically black and white names. The students' infractions were related to insubordination or class disturbance. After each infraction, participants were asked to rate the students' behaviors (scale ranging from 1, not at all to 7, extremely). Teachers were asked to rate how likely it was that they would say that the student is a troublemaker (1, not at all, to 5, extremely). The researcher completed a manipulation check by asking teachers how likely it was that the student's name was that of a black person. They asked teachers to guess the likelihood that the student is from a low-income neighborhood as well as to guess the racial demographics of the school the student attended. Ratings for perceived infraction severity, hindrance, and irritation were combined to create one composite ("feeling troubled").

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

An analysis of variance (ANOVA) revealed a main effect of number of infractions, $F(1, 51)=6.89, p=.011$, and a trending effect of student race, $F(1, 51)=2.72, p=.11$. There was a significant interaction between number of infractions and student race, $F(1,51)=6.09, p=.017$. A paired-samples t-test revealed that teachers felt no more troubled from the first infraction to the second when the student was white $t(28)=0.11, p=.917, d<0.1$. Teachers felt significantly more troubled by the second infraction ($M=5.53, SD=1.04$) than the first infraction ($M=4.40, SD=1.48$) when the student was black, $t(23)=4.06, p<.001, d=1.9$. An ANOVA was conducted on how severely teachers felt the student should be disciplined. A paired samples t- test revealed no significant difference in discipline ratings across infractions for the white student, $t(28)=-0.98, p=.336, d=0.37$. Teachers thought the black student should be disciplined more severely after the second infraction ($M=5.13, SD=0.95$), $t(51)=-3.11, p=.003, d=0.8$, than the white student ($M=3.93, SD=1.67$). The second part of the study measured the extent to which teachers thought the infractions were indicative of a pattern and if the pattern mediated the relationship between race and disciplinary action. The teachers ($n=204$) were asked to indicate if they could imagine suspending the student at some point in the future. The second study found that teachers were significantly more likely to imagine themselves suspending the black student in the future ($M=2.28, SD=0.84$) compared with the white student ($M=2.02, SD=0.85$), $t(188)=-2.07, p=.039, d=0.3$. The more likely teachers were to think the student was black, the more likely they were to imagine themselves suspending that student in the future, $r(188)=.15, p=.036$. This study only discussed the race of the teacher and not his or her gender. Additionally, this study did not discuss the impact of students' skin tone on teacher's responses.

Neal, McCray, Web-Johnson and Bridgest analyzed the effect of African American male movement styles on teacher perception (2003). The study aimed at determining how an African

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

American male's cultural movement styles (walking), would cause teachers to perceive the student's level of aggression and achievement. African American male movement was characterized by a stroll. The authors noted that stylized movement is often referred to as the "cool pose," considered characteristic of many men of this ethnic group. The participants for this study were 136 middle school teachers in a suburban school district in a southwestern state. Participants were told they were participating in a study about middle school students and teachers. An African American student and an Anglo/a-American student performed each walking movement (standard and stroll). Four videotapes were developed. Videotape one depicted an African American student demonstrating a stroll. Videotape two depicted an Anglo-American student demonstrating the same stroll. Videotape three showed an African American student demonstrating a standard walk, and videotape four depicted an Anglo-American student demonstrating the same standard walk. Statistically significant differences were found for movement style, $F = 17.80$, $p = .001$, and ethnicity, $F = 4.53$, $p = .04$. Participants perceived the students with a stroll to be lower in achievement, $M = 1.99$, $SD = 4.97$, than the students with the standard movement style, $M = 1.46$, $SD = 4.64$. Teachers rated the European American student with a stroll, $M = -2.76$, $SD = 4.91$, lower in achievement than the African American student with a stroll, $M = -1.26$, $SD = 5.00$. They rated the African American student demonstrating the standard walk as higher in achievement, $M = 2.62$, $SD = 4.87$, than the European American student demonstrating a standard walk, $M = 0.29$, $SD = 4.16$. Results indicated that there was a statistically significant difference for movement style, $F = 31.32$, $p = 0.001$, but not for ethnicity, $F = 0.63$, $p = 0.43$ when analyzing aggression. Teachers perceived the students with a stroll as higher in aggression, $M = 0.57$, $SD = 6.43$, than the students with the standard movement style, $M = 6.16$, $SD = 5.12$ regardless of ethnicity. When analyzing teachers' perception of student

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

need for special education, a statistically significant difference for the main effect of movement style was found, $F = 14.35$, $p = 0.001$, but not for ethnicity, $F = 0.37$, $p = 0.54$. Regardless of race, teachers perceived the students with a stroll as more likely to need special education services, $M = 2.26$, $SD = 0.64$, than the students with a standard walk, $M = 1.87$, $SD = 0.49$. This article did not discuss if teachers' gender or race plays a role in perception in student movement style or how female movements are perceived.

Pigott and Cowen (2000) examined the effects of teacher race, pupil race, and teacher child racial congruence on teacher ratings of the school adjustment of 445 kindergartens through fifth-grade children from 70 classrooms in 24 racially mixed urban schools in Rochester City School District. Most classrooms yielded 8 child participants: 4 African American and 4 white, with 2 boys and 2 girls per group. The two race groups were closely matched by school, grade level, teacher, and socioeconomic status. Ratings were provided by 26 African American and 44 White teachers, matched by age and years of experience. Children's school adjustment was assessed with the two-part, 38-item Teacher-Child Rating Scale (T-CRS) (Hightower et al., 1986). Part I consisted of three 6-item factor subscales reflecting class problems: Acting Out (e.g., "disruptive in class"), Shy/Anxious (e.g., "shy, timid"), and Learning Problems (e.g., "poor work habits"). Teachers rated these items on a 5-point descriptive scale. These items comprised four 5-item factor subscales: Frustration Tolerance (e.g., "copes well with failure"), Assertive Social Skills (e.g., "expresses ideas willingly"), Task Orientation (e.g., "completes work"), and Peer Social Skills (e.g., "is friendly toward peers"). Four items assessed teacher expectations of children's academic progress (e.g., "will graduate from high school"), using a 4-point scale (1 = does not describe to 4 = describes very well). The total scores ranged from 4 to 16. High scores indicated more positive expectations. Using a Multivariate Analyses of Variance (MANOVA),

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

teachers rated African American children higher on Acting-Out and Learning Problems, and lower on Task Orientation and Frustration Tolerance competencies than white peers. African American children were rated significantly more negatively than white children on five descriptors: boisterous, argumentative, disobedient, irresponsible, and unreliable. For child race, teachers judged that African American children would: (a) have more problems that interfere with their schoolwork, (b) be less likely to graduate from high school, and (c) be more likely to have a difficult future. Teachers rated all children as smarter than their grades indicated. This article did not assess teacher perception of students based on skin tone.

Morris (2007) recognized a gap in the literature about the perceptions and experiences of black girls in classrooms. Using intersectionality as the theoretical framework, this two year ethnographic study was conducted at a middle school. The faculty of the school consisted of approximately 60 educators who ranged from teaching veterans (close to 30 years of experience) to first- year teachers. An African American woman served as the school's principal, and the faculty was roughly two-thirds African American and one-third white, with just a handful of Latina/o and Asian American teachers. The school was located in a predominately low income and working-class area. Using a qualitative methodology (purposive sampling), Morris conducted 14 semi-structured interviews with teachers and administrators. He used a modified version of grounded theory to analyze the data and coded the data using focused coding. He found that the treatment of black girls was related to the influence of dominant ideas of race, class, and gender. Although the black girls enrolled in rigorous courses and participated in class discussions, their actions were viewed as not being "lady-like". Educators at the middle school perceived black girls as challenging authority even though this behavior may be seen as favorable for middle class and white students. Some teachers at the middle school encouraged

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

outspoken behavior from girls, even while describing them as ladies. Most teachers tried to mold these girls into the stereotypical definition of femininity (passive and docile), but instead were creating girls that were less active learners. This article did not discuss if skin tone played a role in teacher perception.

Blake, Butler, Lewis and Darenbourg (2011) studied discipline infractions assigned to black female students enrolled in an urban school district. They explored whether the pattern of discipline infractions and sanctions imposed for black girls disproportionately differed from all female students. Participants were elementary and secondary female students with at least one discipline sanction ($N = 9,364$) enrolled in a midwestern urban school district during the 2005–2006 academic school year. The school district sampled served approximately 32,183 students across approximately 44 schools. School records identified 38 possible discipline sanctions. Sanctions ranged in severity from a warning to expulsion. They assessed black girls disproportionate representation in discipline sanctions using the Relative Risk Ratio (RRR) conducted MANOVA and Descriptive Discriminant Analyses (DDA) to identify the types of infractions that black females experience. The RRR estimates the degree of overrepresentation or underrepresentation of the target group compared to the comparison group. If the RRR is 1, then the risk of the target group is equal to the comparison group. An RRR greater than 1 indicates overrepresentation of the target group. An RRR less than 1 indicates underrepresentation of the target group. RRR Results showed that black female students were overrepresented in all discipline sanctions, $RRR = 2.38$. With respect to exclusionary discipline practices, their results suggested that black girls are twice as likely to receive in school, $RRR = 2.62$, and out-of-school suspensions, $RRR = 2.83$, than their same-gendered peers. A significant main effect for race, $F(36, 32,807) = 16.66, p < .001$, was found indicating that black, Hispanic, and white females

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

differed with respect to the reasons that they were disciplined. Black females were disciplined for defiance, inappropriate dress, using profane language toward a student, and physical aggression. White females were disciplined for truancy. Hispanic females were more likely to be disciplined for a wide range of reasons. Once again this study did not address if skin tone bias was present or if educator's race or gender impacted perception of students.

Hines-Datiri and Carter Andrews (2017) explored black female student high suspension rates when compared to white female counterparts. Authors used Critical Race Feminism (CRF) and figured worlds as the theoretical frameworks to understand the impact zero tolerance policies has on black girls. CRF discusses the experiences women of color face by examining intersectionality of race and gender simultaneously. Figured world framework discusses the context where culture and identity are formed, which in this case would be the classroom. A triangular push-out process is used to describe teachers' perception of black females according to racialized and gendered norms of white students. This process includes using white norms to analyze their identities, actions and determine the type of punishment. Black female students are "... penalized for being their authentic selves" (p, 9, 2016). Black females are historically and socially viewed as not representing white femininity. Black female youth, particularly in the urban school setting, are viewed as having attitudes and being loud. These attributes are used to demonstrate resistance to white femininity that disregards black female youth identity as being unimportant. Zero tolerance school discipline policies cause black female youth to feel as though they are under surveillance and do not experience the "white innocence" that white peers are granted.

Ragsdale (2013) completed an ethnographic dissertation study about three white male assistant principals' perceptions of black male students in a suburban school district. The three

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

administrators were interviewed and observed in the school setting. The study was divided into three stages: initial, follow-up, and final interview and took place in a school district that is composed of approximately 33,000 students. Fifty-seven percent of the students are considered economically disadvantaged, 33% are considered Limited English Proficient (LEP), and 52% are considered at-risk. White males were selected as participants because this demographic are becoming building principals more than any other race and gender (despite the majority of teachers being female). Upon completion of this research, three themes emerged about the types of perceptions white male teachers have of black male students: 1) contingent perceptions, 2) unquestioned perceptions, and 3) unconscious perceptions. A contingent perception is what someone thinks to be true, but recognizes that it not yet certain. In this study a contingent perception was revealed when assistant principals thought black children struggled in schools but could not determine what factors were truly the cause. An unquestioned perception is what someone believes to know is true. During the interviews the assistant principals made statements that they full heartedly believed. They believed that they played a major role in the lives of the students and that all students shared the same struggles as black boys (unmotivated or not interested in education) except that most black males come from single parent homes. All participants noted lack of parental support and role models as a factor contributing to the difficulties black males face in the school system. They acknowledged with no uncertainty that the black males in their building came from different socioeconomic statuses than their white peers. Administrators were asked to describe black males in a few words or phrases. The purpose was to see if they would describe black male students using positive or negative words. One participant responded overly positive (respectful, kind, and fun), one overly negative (attention getting, stuck in a cycle and withdrawn) and the last, balanced (diligent, outgoing, effusive and

vexing). The campus with majority white students had an assistant principal with the most negative descriptors for black male students. The campus with the most students of color had an assistant principal that used the most positive descriptors. The campus that had a balance of white and students of color also had a balance of positive and negative descriptors. It appears then that there could be a link between the frequency of exposure to black male students and perceptions of the assistant principals.

Heilburn, Cornell, and Lovegrove (2015) investigated an association between principal attitudes toward zero tolerance and suspension rates for white and black students in 306 Virginia high schools. Regression analyses showed that principal endorsement of zero tolerance was moderately associated with suspension rates for both white and black students. Paired-samples t-tests showed statistically significant differences in the types of offenses that resulted in suspensions, with Black students significantly more likely to be suspended for disruptive offenses and white students more likely to be suspended for alcohol and drug-related offenses. The sample for this study included 306 high school principals who completed the 2012 Virginia School Safety Audit. The Zero Tolerance Attitudes (ZTA) Scale is a brief measure that was included in the 2012 Safety Audit. The ZTA scale ($\alpha = .75$) was developed from the best discriminating items on zero tolerance from Skiba's 60-item Disciplinary Practices Survey (Skiba & Edl, 2004). Principals were asked to rate their agreement with the following statements on a scale of 1 (Strongly Disagree) to 4 (Strongly Agree): 1) Zero tolerance makes a significant contribution to maintaining order at this school; 2) Zero tolerance sends a clear message to disruptive students about appropriate behaviors in school; 3) Suspension is a necessary tool for maintaining school order; and 4) Schools cannot afford to tolerate students who disrupt the learning environment. Higher scores indicated more favorable attitudes toward the use of zero

tolerance and suspension. The Virginia Department of Education provided school-level discipline data for all incidents that resulted in suspension from the 2011 to 2012 school year. The average Virginia public high school gave 85 total suspensions a year to an average of 47 students per school. At least one suspension was imposed on about one tenth of all students ($M = 11\%$). A dependent-samples t-test indicated significant differences between Black ($M = 17\%$) and White suspension rates ($M = 8.3\%$) within schools, $t(284) = 17.58, p < .001, d = 1.02$. School suspensions were higher in schools where principals supported the idea that zero tolerance disciplinary policies helped maintain order in their schools.

DeMatthews, Cary and Olivarez (2017) used institutional racism, discipline and critical race theory as the theoretical framework to understand how race and school context contributed to the way principals discipline students. Ten secondary principals were selected to participate in the qualitative study based on being in their position for at least two years, working in a school where black students make up 15% of the student body and based on principal diversity. The initial interview included open-ended questions, asking principals about personal and professional background, community and school context, and school and district discipline policies. The second interview asked principals about student race, racism, student discipline, and examples of student misconduct and how they selected an appropriate disciplinary consequence. Afterwards, principals were asked to participate in a focus group and the following themes emerged: social justice dilemmas, priorities, reflection, justifications, and blindness. The principal's ways of handling school discipline issues were classified into three categories: overt racial justifiers (expected students of color and impoverished students to misbehave), rigid rule enforcers (hard facts about the infraction are used to determine discipline consequence) and flexible and cognizant disciplinarians (use a case by case, individualized approach when

selecting a discipline outcome). All principals in the study agreed that students need to adhere to school rules. The concept of equal treatment appeared to be more of a priority than equitable treatment of students during the discipline process. Most principals believed ineffective teacher practices such as lack of teacher/student trust, power struggles, and inconsistent classroom policies perpetuated the discipline gap. DeMatthews et al., (2017) noted that principals are largely at fault for the discipline gap because many do not recognize and confront race-based discrimination that occurs within their schools due to their own bias and implicit bias on the part of teachers and other educational stakeholders. This article discusses how specific principal practices perpetuate the school discipline gap, but did not discuss how specific the population of black students (by skin tone or gender) are disciplined by administrators.

Ferguson (2001) gave a vivid description of educators' perception of black youth in her book entitled *Bad Boys: Public Schools and the Making of Black Masculinity*. She noted that during a district workshop for adult volunteers, the classroom teacher running the workshop described all children as needing attention in order to flourish. The teachers described children as being con artists and manipulators. In her discussion, children become objects of control and are vilified. When children are looked at as the problem, adults become innocent and pure. Ferguson discussed the idea that childhood is defined differently and has changed socially. Childhood means something different depending on location, class, gender, race or a combination of all four. In particular, African American children have been viewed differently because of laws (Brown vs. Board of Education¹⁰), social policy, and media. Black youth are viewed as not childlike and are adultified. She noted, "Adultification is a central mechanism in the interpretive

¹⁰ Board of Education (1954) is a Supreme Court decision that ended the racial segregation of children in public schools violated. The decision noted that segregation violated the equal protection clause of the fourteenth amendment.

framing of gender roles. African American girls are looked at differently than white girls.

African American females are portrayed as mothers, girlfriends, and sexual partners for boys in the room, whereas white females are portrayed as sexually passive and innocent.” (p. 84)

Lindsay and Hart (2017) matched students and teachers based on race in classrooms in North Carolina. The sample included approximately one million students and 50,000 teachers, from the years 2008 through 2013. The purpose of the study was to examine if teacher gender and race impacted the rate at which students receive exclusionary discipline (detention, suspension or expulsion). The data followed students over several years which allowed for comparison of years when a student had a same race teacher. The study found that North Carolina students are less likely to be removed from school as punishment when they and their teachers are the same race. Black students, specifically black boys, are less likely to be subjected to exclusionary discipline when taught by black teachers. They did not find evidence that being matched with the same race teacher benefited white students. Exclusionary discipline rates for black male students were slightly less when they had a black teacher, 13.7% for black female teachers and 12.7% for black male teachers. Black female student's rates of exclusionary discipline was 6.1% for white female teachers, 6.2 % of white male teachers, 4.6% for black female teachers and 5.3% black male teachers. Authors did not note if data gleaned from the study is statistically significant.

Annamma et al. (2019) used CRT and Critical Race Feminism (CRF) as theoretical frameworks to examine black girls' exclusionary discipline outcomes. The researchers examined disciplinary data from a large urban school district to assess racial group differences in office referral reasons and disparities for black girls in out-of-school suspensions, law enforcement referrals, and expulsions. The study first used a quantitative analysis of school discipline data to

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

identify patterns in-office referral reasons and exclusionary discipline outcomes for Black girls. Second, the researchers qualitatively explored statistically significant associations to determine whether black girls' office referral reasons were for objective or subjective behaviors and whether they aligned with dominant narratives about black girls. The cross-sectional data set used in this study included all disciplined female students ($N = 3,628$) in Grades K to 12 who were enrolled in DPS ($n = 183$) during the 2011-2012 school year. The sample of disciplined female students was 57% Latino, 9% White, 29% Black, 1% Asian, 3% multiracial, 0.9% Native American, and 0.1% Pacific Islander. Black girls (49%) were most likely (37%) to have their behavior labeled as disobedient or defiant, followed by multiracial (40%), Latina (36%), and Native (36%) students. Black girls (53%) were significantly more likely than all other girls (50%) to be referred for behavior deemed detrimental. In contrast, white girls (44%) were significantly less likely to be referred to the office for this reason. Black girls were also substantially more likely than other girls to have referrals for third-degree assault (5% vs. 4%). However, black girls were significantly less likely than all other girls to be referred to the office for possession of alcohol (0.6% vs. 1%) or drugs (4% vs. 6%). In contrast, Native American (15%) and white (9%) female students were significantly more likely to be referred to the office for drug possession; whereas, Asian (7%) and white (3%) girls were significantly more likely to experience referrals for alcohol use.

Marucci (2019) looked at seven types of discipline decisions: teachers' likelihood to (a) send the student to the hallway, (b) send the student to a refocus area within the classroom, (c) send the student to a buddy teacher to cool off, (d) write a referral to an administrator, (e) write a referral to a guidance counselor, (f) recommend in-school suspension, and (g) recommend out-of-school suspension. A survey featuring racial priming was administered to teachers ($n = 287$).

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Teachers were randomly assigned to either the African American condition or white condition for the vignette, or what is referred to as the racial prime. Each condition was identical except for one word describing the racial identity of the student. The vignette reads as follows:

“You are watching an educational movie during class. An African American or white student, who consistently arrives late, walks into your room 10 min after the movie starts. He goes to his desk, greeting other students by name as he does so. After 5 min of watching the movie, he begins to throw paper balls into the trashcan that is on the other side of the room. You ask him to stop two times over a period of 5 min. Both times, he looks at you, laughs, and turns his back to you. After the second time, you walk over and take the balls from him. As you are walking away, you think you hear him mutter an expletive under his breath, and the students immediately next to him laugh. He takes out more paper and makes more balls. He continues throwing the balls across the room, and the other students are noticeably getting distracted and start laughing when he misses and cheering when he gets it in.” (p. 55)

After reading the vignette, teachers were asked questions about how they would hypothetically respond to the student in the vignette given either punitive or rehabilitative choices and their overall support for exclusionary discipline as a policy. Teachers responded to series demographic and professional questions. The end of the survey included a recall question on the prime condition and were asked to recall the racial identity of the student in the vignette. The findings showed that white teachers were more likely to ask a white student to sit in the hallway ($M = 4.99$, $SE = 0.21$) than a black student ($M = 4.36$, $SE = 0.19$). These results were significant ($t(257) = 2.21$, $p < 0.05$; $d = 0.28$). White teachers were also significantly more likely to write a referral to an administrator for a white student than a black student ($M_{White} = 5.47$, $SE_{White} = 0.17$; $M_{Black} = 4.27$, $SE_{Black} = 0.18$; $t(256.9) = 4.90$, $p < 0.00$; $d = 0.60$). Similarly,

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

white teachers were more likely to recommend In School Suspension (ISS) for a white student over a black student ($M_{\text{White}} = 3.62$, $SE_{\text{White}} = 0.19$; $M_{\text{Black}} = 2.90$, $SE_{\text{Black}} = 0.15$; $t(238.1) = 2.98$, $p < 0.01$; $d = 0.37$) as well as recommend Out of School Suspension (OSS) for a white student over a black student ($M_{\text{White}} = 1.88$, $SE_{\text{White}} = 0.10$; $M_{\text{Black}} = 1.51$, $SE_{\text{Black}} = 0.08$; $t(230.8) = 2.58$, $p < 0.001$; $d = 0.33$). White teachers were also significantly more likely to support exclusionary discipline if they were in the White condition ($M_{\text{White}} = 3.79$, $SE_{\text{White}} = 0.15$; $M_{\text{Black}} = 3.32$, $SE_{\text{Black}} = 0.13$; $t(257) = 2.44$, $p < 0.01$; $d = 0.30$). Results suggest that teachers were harsher against white students, which is counterintuitive to what we know about anti-black implicit bias and racial disparities in discipline. Marucci (2019) suggests that social desirability is at play in the finding and attempts to override anti-black bias in certain contexts but could lead to a harmful lowering of expectations for black students.

In Summary, the current literature on educators' perceptions of African American students is less than desirable. The research discussed the cultural mismatch that occurs in American public schools because of the lack of teacher diversity. Black students are more likely to be disciplined for subjective infractions. When disciplined, white educators perceive a second infraction by black students as being more troubling than a second infraction by a white student. Teachers perceive black females as more disruptive and less "ladylike" than other race peers. Black students received more favorable ratings from black teachers. The student demographics at a school impacts the way administrators perceive black males. White administrators at predominately white schools view black males negatively, whereas white administrators at predominately black schools view black males positively. Black male movement style also impacted educators' perception. The black male who walked with a stroll was seen as being troubled compared to the black male with a Eurocentric style of walk that was labeled as a high

achiever. While the literature shed light on this important matter, it did not analyze how student skin tone can alter educators perceptions of students. More research is needed to determine if educators demonstrate a skin tone bias that influences and how frequently discipline occurs. Since skin tone bias is prevalent in American society, it is possible that this form of discrimination has permeated the field of education.

Skin Tone Bias

Skin tone bias and colorism can also be defined as racial phenotypicality bias. This term “describes a form of discrimination in which facial features that are associated with a particular racial or ethnic group form a basis for unfair treatment of individuals with certain features (Maddox, 2004). Colorism is one type of racial bias that can occur in different countries. This dissertation will only focus on colorism or skin tone bias in the United States. The literature shows that bias based on skin tone and facial features are not separate entities but are often interwoven. However, this dissertation will add to the current body of literature by better understanding how skin tone alone impacts educators' discipline decisions. At the same time, future research will include facial features as a variable. When people observe people of their own race and other races, phenotypic features may work together to determine perceived racial categorization. However, as later discussed, white participants rate skin color as the most used feature to determine racial categorization (Brown, Dane & Durham, 1998).

Early research examined individuals, particularly young children's skin tone attitudes using dichotomous (black vs. white) stimuli. Arguably, the most well-known of these studies (and most replicated) was the Clark Doll Study of 1947 (Clark & Clark, 1950). The participants in the study were children ages three to seven years old. The children were asked multiple questions to understand their perception of skin color. The doll studies were important because

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

they allowed researchers to gather information about skin color more directly. Overwhelmingly, the children in the Clarks' studies attributed more positive characteristics to the light-skinned doll and more negative attributes to the dark-skinned doll. The findings revealed a statistically significant difference between racial identification and skin color in the ability of dark and light-skin black children to identify themselves correctly ($p < .05$). Eight-five percent of dark-skin black children identified themselves correctly compared to forty-seven percent of light-skin black children. The majority of children assigned positive statements to the white doll and negative statement to the black doll.

Clark and Clark's study was used in helping prove that segregation was harmful for students of color during the Brown vs. Board of Education trial. Farrell and Olsen (1983) reproduced this study to assess black children's skin tone preference after the black activism of the 1960s. This study used children from eleven kindergarten classrooms (151 black five-year-olds). Instead of using dolls, the researcher utilized the Larson-Olden-Farrell Picture Inventory (LOFPI). The LOFPI consists of eight 6-inch photograph cutouts of dark and light-skin boys and girls. Unlike the Clark and Clark study (1947), Farrell and Olsen found the majority of black children were able to identify the doll that look like herself or himself correctly. Only 47% of participants in the Farrell & Olsen study selected a white doll for positive statements compared to 65% in the Clark and Clark study. Negative statements were assigned to the black dolls 36% of the time, whereas in the Clark study, it was 77% of the time.

The Clark Doll Study has been replicated numerous times via YouTube videos, mostly completed without scientific inquiry or formal methodology. However, in 2010, Cooper recreated the study for Cable News Network (CNN), using a different methodology than the original study. Participants in the pilot study were thirty-three children from eight schools (four

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

in the greater New York City area and four in Georgia). Children were either four, five, nine, or ten years old. The schools were described as affluent, majority-white school, impoverished majority-white school, affluent majority African-American school, impoverished majority African-American school. Middle school children were assessed using an 11-item revised version of the Skin Color Opinions and Perceptions Evaluation (SCOPE), a questionnaire comprised of questions related to children's perceptions of their own skin color. Children indicated their skin color choices using the Visual Inventory for Skin Tone Assessment (VISTA) - a commercially produced color bar comprised of ten colors arrayed across the bar from lightest to darkest. Young children's perceptions of, and attitudes toward, skin color, were assessed using a 9-item revised version of the Skin Color Opinions and Perceptions Evaluation (SCOPE). Early childhood participants indicated their skin color choices using the Early Childhood Version of the VISTA (VISTA-ECV). Early childhood participants indicated their skin color choices using the Early Childhood Version of the VISTA (VISTA-ECV). Among children in pre-kindergarten and kindergarten, there was a statistically significant negative correlation between children's scores for positive attitudes and beliefs and their scores for negative attitudes and beliefs ($r = -.76, p < .0001$), and color rejection ($r = -.45, p < .001$), indicating that as one score increased the other score decreased. Children who selected lighter skin tones for the children with positive traits (smart, nice, good, good looking) selected darker skin tones for children with negative traits (dumb, mean, bad ugly) and vice versa. Findings were similar for middle school children ($r = -.56, p < .0001$).

Brown., et al. (1998) developed the first study aimed at examining the processes individuals use to determine the race or ethnicity of another person. Although this research is over two decades old, it laid the foundation for current studies. The study consisted of 75

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

participants. Twenty-five of the participants were male, 48 were female, 48 were white, and 24 were non-white (13 Africans, 9 Asians, 1 Arabian, and 1 Hispanic). Participants were first given five minutes to write down as many names for race and ethnicity (e.g., African, Asian, Caucasian, etc.) as they could recall. Participants were also asked to list facial features that they believed important in determining race. Participants were asked to rate the importance (1 = least, 10 = most important) of nine major facial features (eyes, nose, hair, cheeks, forehead, mouth, skin color, ears, and eyebrows) when deciding on the race of another person. There were five target races: African, Asian, Caucasian, Hispanic, Unspecified (participant needed to guess). The data were analyzed using a 2 (participant gender) x 2 (participant race) x 5 (target) x 9 (feature) ANOVA in which feature was a within-subject variable. No significant participant race or participant gender effects were found. A significant main effect for target was obtained, $F(4, 56) = 2.46, p < .006$. Participants generally rated facial features as more important for Asian targets than for Caucasian targets. The main effect for feature also was significant, $F(8, 56) = 64.49, p < .0001$. Skin color was rated the most important feature, followed by hair, eyes, nose, mouth, cheeks, eyebrows, forehead, and ears.

In a study completed in Scotland, researchers found that white Scottish participants remembered white faces better than black faces after twenty-four-hour intervals (Shepard, Deregowski & Ellis, 1974). The same held for black Rhodesian participants, remembering black faces better after a twenty-four-hour interval. In a follow-up study, Ellis, Deregowski & Shepard wanted to understand the frequency that white and black participants used different facial features when describing faces (1975). Ellis et al., hypothesized that white people looking at white faces would focus on hair color and texture as well as eye color. They noted, "None of these would be particularly discriminating attributes of black faces" (p. 20, 1975). He was not

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

accounting for the genetic variation between blacks. Ellis suggested that black observers viewing black faces would pay attention to other features such as face outline and skin color. Sixteen colored photographs of faces were used, each measuring 66mm by 83mm. There were four white male faces; four white female faces; four black male faces; and four black female faces. The subjects consisted of twelve black males (mean age=18 years old) and twelve white males (all 16 years old). Both samples completed approximately 11 years of formal education. The sixteen stimulus faces were divided into four sets (each set comprised of a different example of a white male, white female, black male, black female face). Each subject was told to imagine that a friend of his was arriving at a nearby station and that he was unable to meet him. Another person that knows the friend has volunteered to go to the station and needed a description in order to recognize the friend. The study showed that whites often describe eye and hair color. The results did not show that blacks use skin color as a facial attribute. Blacks mentioned hair position, eye size, eyebrows, ears, and chin features. The author concluded that paying attention to facial features is a learned process meaning that it is an acquired strategy for analyzing the classes of faces that we are frequently exposed to. Both black and white participants mentioned more features on black faces, although white participants noted redundant features. White participants were less likely to notice distinct characteristics on black faces except dark-skin, kinky hair, and brown eyes. This inability to see specific characteristics may negatively impact a white person's ability to pick out a particular black face from another black face. The black subjects mentioned more features overall than did the white subjects ($F=6.4$, $df=1/22$, $p<.05$). The black faces elicited mention of more features for both groups of participants ($F=4.5$, $df=1/22$, $p<0.5$). There was no significant interaction between race and face and face of the subject ($F<1$). Although this study was completed over forty years ago, the overall findings remain relevant. This study did

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

not discuss how participants remember people of various skin tones. The authors assume their findings can be generalized to all black regardless of skin tone.

Stepanova and Strube (2012) found that people use skin color rather than facial features for racial categorization. The researchers analyzed the factors that define and moderate racial categorization judgments. This study independently manipulated skin color and facial physiognomy to determine their relative weighting in racial categorization. Participants ($N = 250$) judged faces varying on ten levels of facial physiognomy (from Afrocentric to Eurocentric) and ten levels of skin color (from dark to light) under no time constraints, a modest time constraint, and a stringent time constraint. The study found that participants relied on facial physiognomy more when rating faces of light than dark-skin color. Skin color was a more important cue than facial physiognomy under no time constraints. Still, as time constraints became more severe, skin color's importance decreased, yet it remained a more important cue at extreme physiognomy levels.

Smith-McLallen, Johnson, Dovidio, and Pearson (2006), noted, "Research using the Implicit Association Test (IAT) has consistently shown that White participants demonstrate an implicit preference for white, race-related stimuli over similar black stimuli" (p. 46). Although the study is over eight years old, this was the last study completed about implicit race associations for African Americans. Smith-McLallen, et al., investigated whether or not there was a relationship between general evaluative associations with the colors black and white and IAT responses to race-related stimuli such as stereotypical African American and Caucasian names and photographs of African American and Caucasian faces that are commonly used to measure implicit race associations. The study drew three conclusions from the literature review: 1) that people have more negative associations with the color black than with the color white; 2)

that Caucasians tend to associate more negative qualities with African Americans than with whites, and both African Americans and whites tend to associate more negative characteristics with darker-skinned African Americans and; 3) that there is a relationship between preferences for the colors black and white and explicit measures of racial bias.

Hagiwara, Kashy, and Cesario (2012) found that people use phenotypes (visible characteristics) to make inferences about the degree to which stereotypes about the racial group apply to the individual. The participants were 186 self-identified white undergraduates who received partial course credit for their participation. Up to six participants reported to each laboratory session. They were informed that the study examined memory and judgment skills and told that their task was to complete several short computer tasks. The findings showed that whites reacted more negatively toward African Americans with darker skin tone and more prototypical facial features than to African Americans with lighter skin tone and fewer prototypical features. However, the authors noted that Caucasians might be unaware of the negative reactions they have towards African Americans' phenotypes.

Skin tone has emerged as the most allegedly diagnostic feature in racial judgment and is perhaps the most infamous phenotypic feature (Secord, Bevan, & Katz, 1956; Stepanova & Strube, 2009). The article, *When An "Educated" African American Man Becomes Lighter in the Mind's Eye: Evidence for a Skin Tone Memory Bias* (Ben-zeev, Dennehy, Goodrich, Kolarik & Geisler, 2014) noted that the human mind strives for cognitive consistency. In the study, educated African American males tended to be remembered by people as "whiter" in accordance with cultural beliefs, which offers support for the existence of a skin tone memory bias. The authors noted, "These findings, a counter-stereotypic (expectancy-violating) but not a stereotypic (expectancy-congruent) effect-are consistent with the mind's striving for cognitive consistency

or the tendency to attempt to resolve an incompatible cognition in the direction of a stereotype” (Ben-zeev et al., 2014, p. 12). The word “educated” was used as a counter-stereotypic condition, while the words “ignorant” and “athletic” were used for the stereotypic condition. Participants consisted of 125 multi-ethnic and multi-racial undergraduate students. Participants were told that they were participating in a study on memory for faces. Participants were scored based on the number of errors they produced during the recognition memory task for the targets face.

Hannon and DeFina (2016) noted a significant increase in the number of discrimination claims based on skin shade. Due to a rise in this specific type of discrimination, Hannon and Defina examined an increasingly popular skin tone scale (The Massey-Martin scale) that includes a professionally designed color guide to enhance rater consistency. Using logistic regression to analyze data from the National Longitudinal Survey of Youth (NLSY) and General Social Survey (GSS), the authors found that despite adding the color guide, the race of the interviewer matters for the assessment of respondent skin tone. The researchers selected self-identified non-Hispanic African American respondents. From the GSS they selected 263 participants and from the NLSY they selected 1, 797 participants. The interviewer was instructed to memorize the scale so that during the interview, the respondent's skin color could be coded without looking at the chart. Skin tones of participants were recorded into three groups: light, medium, and dark. African American respondents with white interviewers were three times more likely to be classified as dark than those with an African American interviewer. African American interviewers were less likely to assign the highest values of skin darkness to African American respondents than white interviewers (17.8 vs. 34.2% of respondents). Six percent of African American respondents classified by white interviewers were assigned the lowest 3 values of skin darkness as opposed to 18.1% of those classified by African American interviewers. The

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

study found a statistically significant difference in skin tone variation for African American respondents interviewed by African American and white interviewers ($SD=2.0$ for African American interviewers versus $SD=1.79$ for white interviewers, $p<0.01$). In conclusion, the researchers found that other-race bias is linked to the literature on own-race bias. Interviewer's perception of skin tone is based how different the respondent skin tone is from their own.

In American society, children learn about colorism at a very young age. Children in the Clark doll study had negative associations with the dark-skinned doll whereas the light-skin doll was viewed positively. However, studies on colorism have not been conducted within the context of school discipline. In studies about phenotypicality, skin tone was rated as the most important feature in determining a person's race. Black and white people view people from their own race differently. Whites use eye color and hair color when describing one another. Black mentioned more specific features when describing other blacks. The research showed that whites and blacks are better at identifying people of their own race because they have become familiar with analyzing faces that they come in frequent contact with. This could also impact how educators view students from other races. Ragsdale (2010) showed that white male administrators who worked at predominately white schools view black males more negatively than white male administrators who worked at predominately black or racially mixed schools. The literature shed light on the impact of colorism in society but did not discuss its effect on the discipline gap. The dissertation will analyze if and how a student's skin tone impacts school personnel's discipline decisions.

Skin tone and job selection. Harrison and Thomas (2009) examined the influence of skin tone on job selection. They found that light-skinned African Americans were preferred over dark-skinned African Americans Harrison and Thomas revealed, "skin color is more salient and regarded more highly than one's educational background and prior work experience" (p. 134). The study used 240 undergraduate students (68 males and 172 females) from a southeastern university. Students were given course credits for participating. A semi-racially diverse population was used, although 87.5% of the population was white and 72% female. The research study was a 2 (Gender) X 2 (Résumé Level) X 3 (Skin Tone) between-subjects design. Each participant received a packet that contained one résumé with one of the six pictures on it, and they also received one questionnaire. When the participants received the research materials, they were told that the researchers were interested in how strongly an applicant's résumé influences selection decisions, and that they were viewing résumés of applicants interested in a marketing job with a fictitious corporation. Each participant was exposed to only one résumé and picture, and they were not aware that other participants could have the same (or a different) résumé or picture from them. After reviewing the résumé with an attached picture, the participants completed the questionnaire. Participants were asked to rate how strongly they would recommend the candidate, how likely they, themselves, would hire the applicant in the packet that they had received and their demographic information. The study concluded that light-skin African Americans were more likely to get hired for jobs, which explained why light-skin African Americans were more likely to achieve a higher socioeconomic status than dark-skin African Americans.

Job selection is just one facet in American society that is impacted by skin tone bias. It is almost unbelievable that achieving the "American Dream" is more likely to be achieved because

of skin tone and not by merit. The study cited above sheds light on the outcomes for African American students when they eventually enter the workforce.

Skin tone and adoption. Skin tone bias is also present in the adoption process in the United States. Fedders (2010) published a law review noting the high demand for white babies compared to black infants. Because of the high demand and low supply of white infants, private agencies have made adoption fees higher for these infants. This practice is problematic because it perpetuates the notion that whites are superior to other races. More white families enter the adoption process and typically request white babies. Although when white babies are not available, families often accept Latina/o, Asian, Native American, or babies born to one black parent and one parent of a different race. Fedders (2010) surveyed fifty-six private adoption agencies that process domestic adoptions. The results revealed that ten, or approximately eighteen percent of agencies, charge higher fees for the adoption of white infants than black infants, although this practice is not advertised. The Adoption Support Center in Indianapolis, Indiana, lists one price for the adoption of white and Hispanic infants and another, lower price, for infants that are black or mixed with black. Mississippi Children's Home Services advertises that they offer an income-based sliding scale for those individuals who adopt either a black or half-black child. Analysis of the law showed that demanding higher fees for white children is not illegal. Many states subsidize adoptions of children who have special needs or are hard to place. Children can be placed in this category if they are black. The federal tax code allows for a supplemental adoption credit in addition to the standard adoption credit to those parents who adopt a child identified as being special needs. The 2006 Internal Revenue Service (IRS) regulations note that children belonging to a racial minority group can be classified as "special

needs” for purposes of tax credits. This tax credit still exists in 2016 and can be verified by viewing different adoption agencies and IRS websites.

Skin tone and marriage. Hamilton, Goldsmith, and Darity (2009) were propelled to study black marriage rates when they noticed a steady declining marriage rate for black women. To better understand the decline in marriage rates for black women, the authors noted that there are not enough marriageable black men. Black men who are marriage material (employed, not drug/alcohol dependent or not in prison) are highly desirable and have options for spouses. Data from the Multi-City Study of Urban Inequality (MSCUI) was used to estimate the influence of skin shade on the likelihood of ever being married for black women. MCSUI interviewers graded respondents on skin shade, using a Likert scale. The interviewers used three categories: “dark,” “medium,” and “light” to describe the complexion of blacks who participated in the survey. The study found that in young women ages 16-29, 55 percent of light-skinned black females had been married, but only 30 percent of those with medium skin shade and 23 percent of the dark-skinned females had ever been married. Hamilton et al., noted that previous research revealed a higher social status for light-skinned black women. Marriageable black males may prefer a high status (or light-skinned) spouse to gain acceptance in certain social circles and gain promotions at work.

Skin tone and classroom experience. Hunter (2016) acknowledged that classroom experiences vary for African American and Latina/o students. Specifically, her theoretical work focused on understanding the social practices in schools that create differences in student experiences based on skin tone. Hunter’s research discussed the classroom-level interactions between students, teachers, parents, and administrators. Hunter made many valid points that are rarely considered. Hunter acknowledged that racialized beauty norms impact students,

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

specifically females. Lighter skinned females are viewed as more aesthetically appealing because they favor whites more so than dark-skinned females. This beauty queue impacts students' involvement in homecoming queen elections, sports teams, and student government organizations. The halo effect plays out in the classroom when lighter-skinned students of color may seem more attractive to teachers because of racialized beauty standards that favor European features (light hair, light eyes, or other Anglo features). Light-skinned African American and Latina/o parents have racial capital in their child's school, which can be transferred into economic capital and appear to have more status in the eyes of white teachers and administrators.

Hall's (2016) research explained how African American and Latina/o students are impacted by the Bleaching Syndrome during their educational careers. The Bleaching Syndrome impacts perceptual attributes, behavioral traits, and causes people to ascribe traits to a person based on colorist stereotypes. The Bleaching Syndrome is a theoretical construct conceptualized by Hall (1994) and refers to a gradual erosion of the quality of life and human well-being based on racial and colorist forces. The construct is initiated by the perceptual, followed by the psychological, and concluded via the behavioral. Those impacted by the bleaching syndrome begin to subconsciously internalize European ideals and adopt specific behavioral traits that may conflict with their own inner beliefs. Hall provided an example of an excerpt from *The Crisis*, the official publication of the NAACP, W. E. B. Du Bois, a light-skinned African American, labeled black leader Marcus Garvey as "fat, Black, and ugly," implying that dark features were unattractive (Lewis, 2000, p. 80). According to Hall, educators play a major role in dismantling colorism and the bleaching syndrome. He concludes by providing strategies for educators and schools. These strategies include making sure school personnel are aware of colorism by including it in the staff handbooks, teaching students how to constructively negotiate experiences

dealing with colorism such as name-calling, and developing an awareness of how colorism causes student profiling, stereotyping, and educational outcomes.

Skin tone and educational achievement. Using data from the National Longitudinal Study of Adolescent Health and Adolescent Health and Academic Achievement (AHAA), Thompson and McDonald (2015) showed significant skin-tone differences in grade point average (GPA) both across and within racial groups. The study included Asian, black, Hispanic, and white students. The researchers used three waves of the Add Health survey data (representative of grades 7 through 12 in the United States in 1994) to understand the extent to which the self-identified race category and perceived skin tone influence educational performance independently and simultaneously. The data was collected through in-depth interviews with adolescents and their parents. An in-school survey (1994) and three waves of in-home interviews (1995, 1996, and 2001) were completed. The child's skin tone was measured by the interviewer. The study highlighted two important findings. First, the study found that dark-skin-tone students' GPAs were .4 points lower than light-skin-tone students ($-.080 \times 5 = -.4$) in all racial groups. The gap in GPA between dark-skinned and light-skinned students could not be explained by family characteristics, student-teacher interactions, or school context. However, skin tone differences were not associated with a decline in achievement for African American students. For Asian, Hispanic, and white students having a darker skin tone was associated with having a lower GPA. Their findings suggest that African Americans, regardless of skin tone, have lower GPAs than their other race peers. Secondly, the data showed that teacher bias and exclusionary rules contribute to the skin-tone gap in educational achievement. However, their findings suggest that those factors only explain a relatively small portion of that variation.

Skin tone and teacher preparation. McGee, Alvarez, and Milner (2015) analyzed the experiences African American female students endure concerning their skin color, outside of and within school environments. The researcher interviewed 36 high-achieving high school female students. In this study, 8 of 16 teenage girls self-identified as brown-skinned or dark-skinned. Out of those eight, seven characterized their educational experiences as stressful, in part because of their skin color, as they noted that their physical features were devalued and linked to negative stereotypes and perceptions. These students expressed the desire to have lighter skin because students with lighter skin appeared to benefit from heightened academic and social advantages. According to the group of brown/dark-skinned students, lighter-skinned black female students at their schools were often prom queens, in school leadership positions, and were sought after for romantic relationships. Lighter-skinned black females benefited from teachers' positive perceptions of their academic abilities. Dark-skinned students, females, in particular, experience skin color dissonance. At a young age, children observe the different treatment of peers based on skin tone, and often it is positive treatment of those with lighter skin. McGee et. al., noted that students of all races already have to deal with the stress of succeeding in the K-12 setting. Race, specifically skin tone, can be an additional stressor for dark-skinned students. Dark-skinned students, female in particular, experience skin color dissonance. At a young age, children observe the different treatment of peers based on skin tone, and often it is positive treatment of those with lighter skin. McGee et. al. noted that students of all races already have to deal with the stress of succeeding in the K-12 setting. However, race and skin tone can be an additional stressor for dark-skinned students. McGee et al., (2015) noted, "As darker-skinned youth move toward adolescence and become more cognizant and sensitive about their appearances and those of others, it can be a troubling task to develop a stable educational experience when distracted by

the onus of discrimination correlated to their color.” (p. 73). Teacher education programs should include learning opportunities on the construction of race as a phenotype (the physical construction of skin tone, hair texture, facial features, and body physique) as an influence on the thinking, beliefs, and consequent practices of teachers in P–12 classrooms. Programs need to address the history behind colorism in education and its prevalence in all areas of society. The concept of colorism is a complex topic, which involves critical thinking and questioning.

Kunesh and Noltemeyer (2015) used an attribution-based theory of stereotypes as the theoretical framework in a qualitative study about beliefs pre-service teachers hold about black boys. Attribution based theory of stereotypes states that under stress, people form judgments of others to make quick yet biased opinions using observable characteristics (age, race, weight, etc.). The authors linked this theory to the quick and sometimes stressful decisions classroom teachers must when disciplining students. Participants in the study were 148 undergraduate education majors or pre-service teachers. Participants were randomly divided into three, almost even groups. Each group received an email inviting them to participate in an online questionnaire for a chance to win a gift card. Participants were randomly assigned to one of three surveys. Each scenario was about a seventh-grade male student. The child's gender and grade were specified. In the situation, the student is not completing the assigned task, and he responds insubordinately to the teacher's prompt to begin working. No or little information is given about the student's background, the student-teacher relationship, or the nature of the act (tone of voice). Names were used to imply race. In the first vignette, the name Darius (stereotypically black) was used. The student in the second vignette was named Cody (stereotypically white). The first and second vignette also explicitly said the race of the male student. The third vignette used the stereotypical black name but did not explicitly state the race of the student. Participants

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

who were randomly assigned to read the vignette featuring a Black student whose race was implied through the stereotypical name alone had 6.187 times higher odds ($p = 0.003$) of reporting that recurrence was “likely,” compared with participants who were randomly assigned to read the vignette featuring a White student. Participants who were randomly assigned to read the vignette featuring a Black student whose race was explicitly stated had 3.962 times higher odds ($p = 0.013$) that recurrence was “likely,” compared with participants who were randomly assigned to read the vignette featuring a White student. Results showed that educators might be engaging in implicit bias and discrimination.

Skin tone and school suspension. Research has shown a link between skin tone and school suspension. Hannon, DeFina and Bruch (2013) used data from the 1997 National Longitudinal Survey of Youth and the National Longitudinal Study of Adolescent to Adult Health (Add Health). They found that having a darker skin tone significantly increased the odds of suspension for African American adolescents. African American females with the darkest skin tone were statistically significantly three times more likely to be suspended than the lightest skinned African American females. The findings remained consistent even when using similar measures in a different sample of African Americans from the national study. African American male student suspension probability varies from 48% when skin is lightest to about 69% when skin is darkest, which is a 21% rise. For African American females, the probability doubles, with a probability of suspension rising from 28% for lightest skin tone to 58% for darkest skin tone. The authors concluded, “The results suggest that discrimination in school discipline goes beyond the broad categories of race to include additional distinctions in skin tone. Skin tone bias is “the tendency to judge a target’s membership in racial categories/subcategories and to form behavioral intentions toward the target” (Zeev, Dennehy, Goodrich, Kolarik & Geislert, 2014, p.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

1). Many studies have shown that people have an automatic preference for white Americans.

“One of the evocative characteristics of a skin tone bias is that it is pervasive across and within diverse ethnic and racial groups, including whites, Latinas/os, and African Americans” (Maddox, 2004).

Blake, Keith, Luo, Le, and Salter (2017) examined the role of colorism in explaining the disproportionate discipline experiences of adolescent African American girls. Using data from Waves 1 and 2 of the Add Health data set, a nationally representative sample of adolescents (seventh to 12th grade) first surveyed in 1994–1995 and followed through adulthood, the researchers used the data sets to determine school suspension rate, behavioral functioning, academic functioning and school climate and skin tone/race-ethnicity. First, students' self-reported ethnic/racial background was coded as either white, African American, or Latina/o.

African American and Latina/o students' skin tones were then rated on a 4-point scale by the raters. African American skin tone data was used because of the lack of variability in rating by interviewers for Latina/o students. African American students were coded as either dark-skin, dark-brown skin, medium-brown skin, and light brown/white skin. White students served as the reference group. When controlling for school's use of punitive discipline policies, students' prior discipline history, behavioral and academic functioning, and perceptions of school climate, colorism was a significant predictor of African American females' risk for school suspension. Using a logistic regression analysis with odds ratio, African American females with the darkest skin tone ($p=.002$) were twice as likely to be suspended as white females but this

relationship did not hold true for light-skin African American females ($p=.885$) or medium-skin African American females ($p=.114$).

These two studies, to the best of my knowledge, are the only studies examining student skin tone and school discipline rates. There are many questions dealing with colorism and school discipline that are unanswered. This study did not obtain information about the school or teacher demographics of the school the student attended. Additionally, Hannon et al., (2013) and Blake et al (2017) used a preexisting national data set using information about children who were ages 12-16 as of December 31, 1986, before "post-racial" America. While this study has added to the literature on the existence of colorism in education, the data is over thirty years old. Furthermore, participants' skin tones were assessed by the interviewers who did not have any formal training on assessing skin tone. Starting in the second round, interviewers recorded race and gender for each respondent based on the interviewers' observations. In round twelve, the interviewers recorded the color of the respondent's skin using a color card with skin color gradients from 1-10 (1 being lightest and 10 being darkest) to determine a color that most closely corresponded to the respondent's skin tone. My dissertation will provide current data on how skin tone impacts discipline rates.

Skin tone and transitioning. Ryabov (2013) focused on the impact of skin tone on school-to-work and school-to-college transitions of African American youths. The authors used complexion advantage and social mobility as its theoretical framework. The author recognized that previous research has shown that lighter-skinned African Americans are given privileges in society dating back to slavery and have more social mobility (better employment and higher educational levels). The author selected participants from the National Longitudinal Study of Adolescent Health survey. The survey targeted adolescents and was collected in three waves, in

1994–1995, 1996 and 2001–2002. The sample consisted of adolescents who claimed to be African American in Wave 3 (final $N = 2,546$). Skin tone was assessed by the interviewer and rated as either black, dark brown, medium brown, and light brown. Among unemployed African American males, those with black skin tone constituted the largest group (39%). Individuals with medium brown tones were the next largest group among the unemployed (33%). Youths with black skin tone were less likely than youths with dark brown skin to go to college. The odds of having a college education were higher for adolescents with lighter skin tones (i.e., medium brown and light brown shades). Adolescents with medium and light-skin tones were more likely than dark brown-skinned adolescents to have a fulltime job by Wave 3.

Skin tone and in the United States penal system. Although research about school discipline and colorism is scarce, the most closely related topic is colorism in the United States prison system. The literature about colorism in the penal system is included in this literature review because school discipline and prisons share similarities except for different populations (children versus adults). Viglione, Hannon, and DeFina (2011) examined whether perceived skin tone (assessed by correctional officers) of 12,158 black women inmates imprisoned in North Carolina between 1995 and 2009 is related to maximum prison sentence and actual time served. Using publicly available records of the most recent incarceration experiences of black females, the data set contained an indicator variable for whether or not an individual is assessed by a correctional officer to have a light-skin complexion at the time of admissions to prison. Prisoners were coded 0 for non-light and 1 for light-skin. The authors noted that the identification of prisoner's skin tone by intake officers is a common practice in many state prisons across the United States and that there is not a systematic assessment for skin tone. The large sample spanned over many years and therefore included a diverse group of officers with different

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

backgrounds, experiences, and biases. The researchers calculated the maximum time an offender could legally be incarcerated. They excluded inmates that were still incarcerated or serving life without parole. The coefficient for the variable light-skin was negative and statistically significant at the .05 levels indicating that those with light-skin receive a lenient prison sentence ($b = -0.124$). Light-skin women receive approximately 12 percent less time behind bars than their darker-skinned counterparts.

Branigan, Wildeman, Freese, and Kiefe used data from the Coronary Artery Risk Development in Young Adults Study (CARDIA), to determine if black men's likelihood of getting arrested increased with skin darkness as compared to white men of varying skin tones (2016). Respondents were randomly selected after stratification by race, sex, age, and education in four US cities: Birmingham, AL, Chicago, IL, Minneapolis, MN, and Oakland, CA. The skin color measure in CARDIA was measured as the percent of light reflected off the skin, assessed using a Photovolt 577 spectrophotometer at the underside of the upper arm. A logistic regression formula was used to determine if skin tone does predict the likelihood of arrest among White men ($p < 0.01$), while for Black men, there is no significant association. The authors conclude that whites are often thought of as lacking skin tone, and research often views all whites as having the same degree of lightness. However, in this study, the darker the white male, the greater the probability of arrest. It is important to note; the darkest-skinned white person is still less likely to be arrested than the lightest-skinned black person. Black men's probability of arrest and incarceration remains constant across the spectrum of skin color in comparison to white men. The interracial disparity between black men and the likelihood of arrest.

Gyimah-Brempong and Price (2006) noted that African American male inmates that have a lighter skin tone received shorter prison sentences. The authors analyzed the effects of skin hue

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

on the crime hazard and prison sentence using data on African American offenders with last names starting with the letter "A" incarcerated in the state of Mississippi as of August 20, 2005.

The authors concluded that being black and having a dark-skin hue induces a transition into criminal activity by increasing the probability that criminal activity is acceptable because of a decreased income compared to lighter skin African Americans. In other words, African Americans with a darker hue are more likely to engage in criminal activity because of decreased opportunities to generate income due to skin tone.

Skin tone bias also has an impact on the amount of time African American inmates spend in jail. Individuals with more stereotypically African American features (dark-skin, full lips, coarse hair, and broad nose) were more likely to be falsely identified as criminals (Eberhardt, Davies, Purdie-Vaughns, and Johnson, 2006). Eberhart et al., examined whether the likelihood of being sentenced to death is influenced by the degree to which a black defendant is perceived to have a stereotypically black appearance. The authors used a database containing over 600 death-eligible cases from Philadelphia, Pennsylvania, from 1979 and 1999. Forty-four of these cases involved black male defendants who were convicted of murdering white victims. Photographs of the black defendants were presented in a slide-show format (4 seconds per slide) to undergraduate raters who did not know that the pictures depicted convicted murderers. Raters were asked to rate the stereotypicality of each black defendant's appearance. They were told that they could use features (lips, nose, hair texture, skin tone) rate, defendants. The results showed that defendants whose appearance was perceived as more stereotypically black were more likely to receive a death sentence than defendants whose appearance was perceived as less stereotypically black, $F(1,36)= 4.11, p<.05$). Perceived stereotypicality of black defendants convicted of murdering black victims did not predict death sentencing, $F(1, 110) < 1$.

The research discussed in this section is important because data about colorism in the adult criminal justice system closely resembles the intraracial discipline gap in American public schools. Earlier in this literature review, the adultification of black children was discussed. The literature shows that African American children are perceived as having adult like qualities by educators. The literature about colorism in the penal system mirrors the findings of Hannon, Defina and Burch (2013), except that one study assessed school age children while the other used adult participants. The fact that the outcome for dark-skinned African Americans is parallel within the area of discipline, regardless of age, signifies that colorism is an issue in society that has filtered into the public school system. Dark-skin black children may be treated like dark-skin black adults because of skin tone bias.

Skin tone and perceived intelligence. Hannon analyzes colorism through the lens of white Americans (2015). He noted that colorism is typically framed as intraracial (black-on-black) discrimination. Hannon used data from the 2012 American National Election Study to analyze white interviewers' perceptions of minority respondent skin tone and intelligence ($N=223$). Interviewers evaluated respondent's' apparent intelligence on a 5-point scale coded as (1) very low, (2) fairly low, (3) average, (4) fairly high, and (5) very high. Respondents covered the entire spectrum of possible skin tones (1-10) as defined by the Massey-Martin Scale (Figure 1.1)

Results from ordinal logistic regression analyses indicate that African American and Latina/o respondents with the lightest skin are several times more likely to be seen by whites as intelligent compared with those with the darkest skin. Using ordinal logistic regression, the researchers found that skin darkness was highly statistically significant ($p<.001$). Using the data, they concluded that a white observer would look at two identically qualified minorities and

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

assess the lighter-skinned as more intelligent. A one standard deviation increase in skin lightness roughly triples the probability of being perceived as having above-average intelligence. The authors concluded that if white adults have a tendency to equate lighter skin with intelligence, then the quality and level of expectations, white teachers, and other school authorities have for certain students may be impacted. He noted that more research is needed to understand how stereotypes based on skin tone can create self-fulfilling prophecies in educational achievement and school disciplinary actions. Colorism is not monitored or mentioned enough in any field of study. Educational institutions should track student outcomes by race, ethnicity, and skin tone to better understand disproportionality outcomes.

This research is valuable because this phenomenon could also be seen in American public schools. For example, although not revealed in the study, some of the participants could have been schoolteachers. If beliefs about intellectual inferiority, based on skin tone, are prevalent in the larger American society, then it is highly likely that these beliefs have permeated American public schools. Although this dissertation did not analyze the educator's perception of students' intellectual capability, it does highlight what current literature on colorism has proven and can aide in the interpretation of data.

Own-Race Bias

Hills and Lewis used the term own-race bias (ORB) to describe the “failure to generalize expert perceptual encoding developed for own-race faces to other-race faces” (p. 996, 2006). ORB can also be described as the other-race effect, the own-race effect, the other-ethnicity effect, and the cross-race effect (Bukach et al., 2012 & Walker and Hewstone, 2006). It ties into

the idea that people of another race "all look alike"¹¹. This research was used when five Black men (Quincy Five) were wrongfully charged with murder because five white eyewitnesses positively identified them as the killers. After forensic evidence was obtained, the five men were exonerated, and three other men were convicted. Meissener and Brigham (2001) completed a meta-analysis and empirically reviewed over 30 years of research on the ORB in memory for faces. Data was analyzed from 39 research articles, involving 91 independent samples and nearly 5,000 participants. Fifty-six percent of the samples were reported white, 32% black, and the other 12% were of Arab/Turkish, Asian, and Hispanic origin. The researchers found that participants of all races were 2.23 times more likely to recognize an own-race face than an other-race face in a recognition experiment. The study aimed at understanding the ORB effect for measures of hits and false alarms. Hits were described as correctly identifying a face as "old" and false alarms as incorrectly identifying a face as "old". Measures of hit and false alarm rates were examined, including an analysis of 8 study moderators. Results indicated a "mirror effect" pattern in which own-race faces yielded a higher proportion of hits and a lower proportion of false alarms compared with other-race faces.

A person of one race is naturally, almost instinctively and unknowingly, better at recalling details about a person from their own race versus another race. The results of the study show that ORB is not related to any level or degree of racial prejudice. Own-race bias offers insight on how people view others from the same and different races. False alarms (incorrect eye witness identification) can be detrimental when determining the innocence or guilt of a person being convicted of a crime based on eyewitnesses' recall of an assailant phenotypic features.

¹¹ This phenomenon was first investigated in a dissertation study (Berger, 1969). Research from the dissertation study was used in the Quincy Five trial (1971) to prove that white eyewitness misidentified five Black men due to the fact that when recalling identifying features of a person of another race, people made generalization and assumed all people of that race all look alike.

Summary

The research presented in this literature review suggests that a subtler form of racism is at play in American society and public schools. Members of society may be under the impression that we are living in a “post-racial” America, citing President Barack Obama’s eight-year tenure in the White House as evidence. In the case of the American schoolhouse, the literature suggests that dark-skin black students are still relegated to second-class citizenship. In the “post-racial” society, integrated classrooms have caused educators to believe that all students have equal access to education. The limited research completed on skin tone shows that dark-skinned students are being disciplined at higher rates than lighter skin blacks and whites, which creates equity issues within a specific population of historically underachieving students.

The literature shows that the interviewer rating of skin tone is not always reliable or accurate. Skin tone ratings are subjective unless a specific guide or scale is used. As seen in the murder trial for Khomas Revels, white people remember people of other races differently and sometimes not at all. Additionally, whites and blacks recall and describe different features of one another. For instance, blacks recount specific features and characteristics to describe other blacks whereas, whites use more general features when describing blacks.

Purpose

The purpose of this dissertation is to increase the understanding of how colorism plays out in the public-school setting, specifically in the context of the discipline gap. I found only one article that has investigated colorism and the discipline gap. Hannon, Defina, and Burch (2013) initiated the discussion of colorism impacting the discipline rate. The study analyzed data from 1996, and interviewers assessed participants' skin tone. More research is needed to understand the discipline gap, not as a gap between blacks and whites, but as a gap that may also be based on colorism. This dissertation continues the discussion of colorism and the discipline gap by updating the methodology. Monroe (2013) recognized this need when she said:

Situating educational phenomena within the interlocking frameworks of racial constructions, racism, and color bias will place theories, policies, and databased patterns in the appropriate sociohistorical context, help unravel the complexity of enduring problems and disrupt simplistic perceptions that diminish the influence of the color complex. (p. 9)

The literature shows that school discipline is widely defined by suspension and expulsion data. This dissertation helps define school discipline by understanding the discipline process in its entirety. This study analyzes the school discipline pathway. Understanding the discipline pathway begins with understanding factors educators use when deciding to reprimand or suspend a student versus only calling a student's parents. This study also makes strides towards understanding if stereotype threat is at play in public school classrooms. Educators may unknowingly hold stereotypes about students and therefore expect a student to behave in a particular manner. Holding this belief about a student causes the child to behave in a way that confirms the stereotype. Few studies investigate the complexity of colorism to examine how this social construct plays out in the classroom. Much of the research on this topic utilizes the same and very common research articles. Researchers analyze the same research as other researchers,

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

however they use a different theoretical framework. Very few researchers are conducting new studies or generating new data, which would aid in better understanding the school discipline gap and its relationship to colorism. This dissertation examines both race and gender to understand how they intersect and unfold in the educational setting and if colorism is at play during the disciplining process.

Chapter 3: Methodology

There are unanswered questions about why there is a disproportionate rate of African American students disciplined in American public schools each year and even more unanswered questions about the intra-racial gap between African American students. Previous research on colorism has shown that African Americans with darker skin have less favorable outcomes in marriage, employment, and prison sentences. While we know that African Americans encounter a variety of bias and discrimination in many aspects of American society, limited research has been completed to understand if school personnel bias, as it pertains to colorism, may contribute to the school discipline disparity. Colorism has been studied in the fields of criminal justice, sociology, and psychology. The school discipline gap has been researched according to race, gender, and student disability (special education). However, the field of education has not thoroughly explored the impact student skin color could potentially have on the disciplinary decisions.

The purpose of this dissertation is to increase the understanding of how colorism plays out in the public-school setting, specifically in the context of the discipline gap. Therefore, this study is designed to answer the following questions:

1. How "harshly" do educators discipline based on scenario type?
2. Are there differences in educator discipline decisions by color of vignette student skin tone?
3. Are there differences in educator discipline decisions by gender of vignette student?
4. Are there differences in educator discipline decisions by intersection of gender and skintone of vignette student?
5. Are there differences in educator discipline decisions by race of educator?

6. Are there differences in educator discipline decisions by race of educator and skin tone of vignette student?
7. Are there differences in the discipline decisions made by black and white educators based upon brown skin tones of light, medium, and dark?
8. Are there differences in educator discipline decisions by race of educator and gender of vignette student?
9. Are there difference in educator discipline decision by gender of educator?
10. Are there differences in educator discipline decisions by gender of the educator and vignette student skin tone?
11. Are there differences in educator discipline decisions by gender of the educator and gender of vignette student?

Methodological Approach

This investigation is an experimental vignette study using survey methods to collect data. The study is aimed at understanding the implicit attitudes of school personnel as they make discipline decisions toward African American students of various skin tones. The treatment condition in this study is student skin tone. The control condition is the discipline scenario without the student pictures. Student gender and educator gender and race are independent variables, and the outcome or dependent variable is educator discipline decision.

Population and Sampling

Participants in this study are former students in the teacher education Master's degree program, educational leadership Master's, Post-Master's, and Doctoral programs at a university in the southeastern United States who were enrolled from 2007-2018. Student names and university and/or alternate email addresses were obtained from the University's Office of Assessment. A total of 1,016 students emails addresses were available. Additionally, alternate email addresses were available for sixty-two participants who provided an email address when applying to a program. Internet search engines were utilized to

obtain alternate email addresses for 195 participants.

Participants were randomly assigned to treatment or control groups. Using an online randomizer, (randomizer.org) numbers were assigned to each email address (range of 1-9). The number the randomizer assigned determined the survey the participant received. To increase the likelihood of rejecting the null hypothesis, and to increase statistical power in the study, a priori power analysis was completed indicating each treatment vignette will need 64 completed responses to meet power requirements. In hopes, of yielding a high response rate, the entire sample was utilized. However, 228 participants completed the study. Survey results were analyzed in the aggregate and not individually to maintain anonymity.

Variables

Treatment Variable

The treatment variable in this study is student behavior presented in a vignette under 9 conditions: 8 vignettes contain a picture of the student in one of 8 gender-skin tone combinations (four skin tones and two genders) and a ninth vignette with no pictures. Although the scenarios are based on real infractions that occurred in the school setting, the wording was carefully selected to reduce bias as much as possible. Scenarios were written to not elicit a specific response, but to allow respondents to naturally respond to the scenarios as if it were playing out in the actual classroom setting.

Scenarios were developed after carefully analyzing the Virginia Department of Education (VDOE) Discipline, Crime and Violence Reports (2013-2014, 2014-2015 and 2015-2016) and in conjunction with information gained from the pilot study. These reports were used to identify the most frequently cited school infractions committed in the state of Virginia over the last three years. The report noted that the top three infractions are: defiance/ refuse request, classroom or campus disruption, and disruptive demonstrations. All three infractions are subcategories of the

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

offense of disruptive behaviors and were ranked in the top five reasons for short and long-term suspension. Infractions falling under the category of disruptive behaviors are defined as "Unwillingness to submit to authority or refusal to respond to a reasonable request. Any behavior that substantially disrupts the orderly learning environment" (Comprehensive User Guide for Discipline, Crime, and Violence (DCV) Data Collection and Submission, 2017, p. 48).

The discipline scenarios were initially created by holding two focus group sessions with groups of educators who work in the southeast region of the United States. The first group consisted of seven K-12 teachers and the second included six 9th-12th grade teachers and two high school administrators. During the sessions, educators spoke about classroom discipline issues that they have experienced throughout the years. Group members' experiences ranged from 5-35 years of teaching or administering in the public school setting. Discipline experiences among the two groups fell within the same infraction categories although specific scenarios were different due to varying student ages. A final group of three administrators were consulted to develop additional scenarios. Final scenarios were selected based on the following criteria: falling under one of the VDOEs top infractions, generalizability across various school districts, and generalizability across current position held in a school division (teacher, support staff or administrator). Additionally, scenarios were critiqued to determine if scenarios were too "boy-like" or "girl-like" instead of gender-neutral. The Disruptive Demonstration and Classroom or Campus Disruption scenarios were believed to contain wording that may not be gender neutral. Using a state-level web-based reporting tool, analysis of suspension data showed that both male and female students are being suspended for the same top three infractions, although males are being suspended at a much higher rate. However, the data does not list the specific details of each infraction to determine the type of actions male and female

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

students display that is classified under the same offense.

A survey was sent to students enrolled in Master's in educational leadership program (separate from participants in the final study) asking them to rate how likely the student in the scenario was a male versus a female student and how the scenarios could be changed to more boy or girl like. Sixteen participants responded to the survey. Over eighty percent of participants stated that student in the disruptive demonstration scenario was likely or very likely to be a male student, while over 60% of participants stated that it was unlikely that the student was female. Suggestion for making the scenario more relevant to female students included: omitting farting sounds, including cell phone usage and highlighting exaggerated body language (eye roll) and talking back to teachers. When analyzing the Campus Disruption scenario, over sixty percent of participants stated that the student was male, while fifty percent of participants reported the student neither likely or unlikely to be female. Suggestion for making the scenario more relevant to female students included having the student pass notes or doodle on paper instead of making paper airplanes, using a different expletive, playing with makeup. Scenarios were further revised upon receiving feedback.

Validity of treatment variable

Student pictures were initially created using the software Make Me Human (see Figure 3.1).

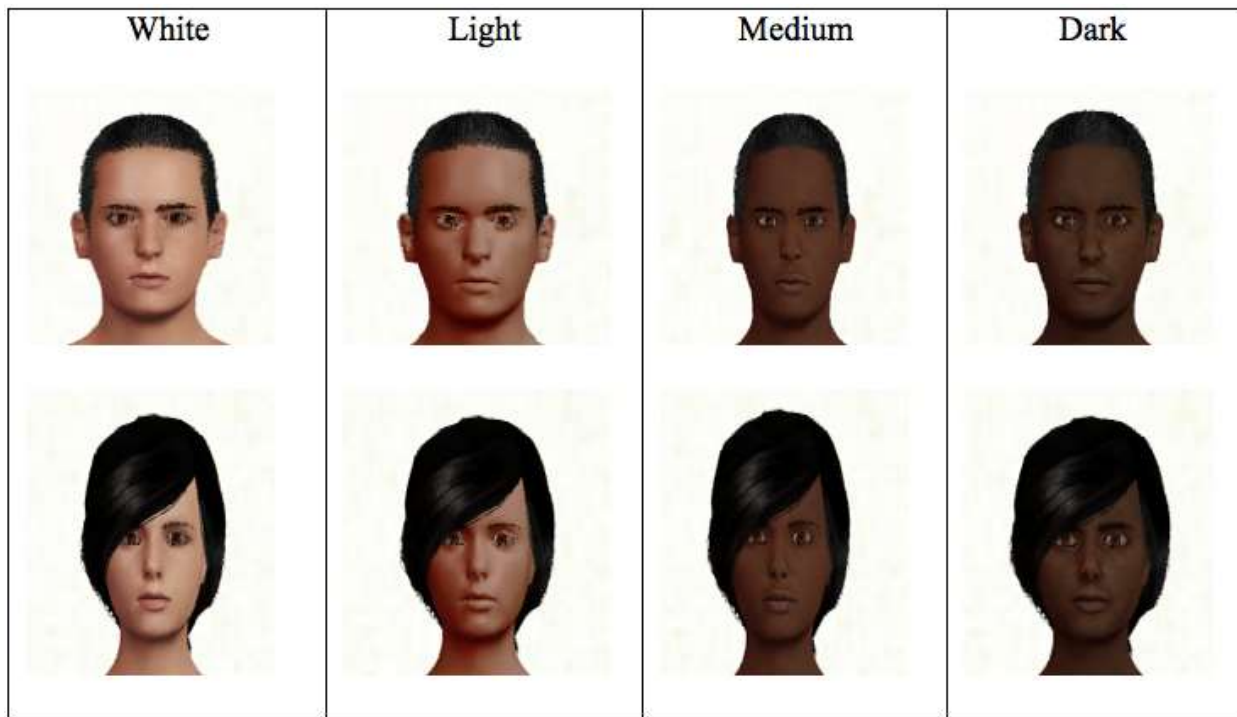


Figure 3.1 Original Student Picture Matrix

Racial identification of student pictures were rated by ten K-12 educators who were chosen through convenience sampling. Educators were asked to identify the perceived race of the student in the picture, “What is this student’s race?” Four participants were asked to rate pictures of the student on paper while the remaining six were asked to rate the pictures via the internet. Although the white, light, medium and dark-skin tones of male and female students were identical, skin tone was not perceived to be the same across gender. Sixty percent of respondents rated the white male student as white and forty percent as Hispanic. Eighty percent of

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

respondents perceived the white female as white, ten percent as Hispanic and ten percent as Asian. Seventy percent of respondents perceived the light-skin male student to be black and thirty percent perceived the student as Hispanic. Racial perception of the light-skin female student portrait varied with seventy percent of respondents rating the student as Hispanic, ten percent said Asian, ten percent white, and ten percent Native American. Ninety percent of respondents perceived the medium and dark-skin tone male and female portrait to be black and ten percent Hispanic.

One respondent noted, "Some of these pictures are difficult to define. Some are ambiguous". Another respondent explained "I am trying to use features such as hair texture and facial features, but that's not helpful". The results from the treatment variable validation study showed that using skin tone to determine race is subjective and in the absence of multiple phenotypic features, skin tone, is used for racial identification. White and light-skin student portraits were noted as being more racially diverse (more than one race or ambiguous). The darker the student skin tone, the more likely they were to be perceived as being the only black. The likelihood of the dark-skin and medium skin student being perceived as Hispanic or Asian decreased when compared to white and light-skin student portraits. This finding, no fault to the student pictures appears to be based on the stereotype held by society that people of a darker hue are not racially diverse and are primarily or solely of African Descent or black. This belief does not consider the various skin tones seen in people of African, Asian and Hispanic descent.

Feedback about the student pictures were considered, and revisions were made based on feedback from respondents and dissertation committee members. A new set of student pictures was developed, using pictures of two real young adults to create photographs that were realistic.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Redeveloped student pictures were obtained through google images and skin tones were modified using Adobe Photoshop software (see Figure 3.2).

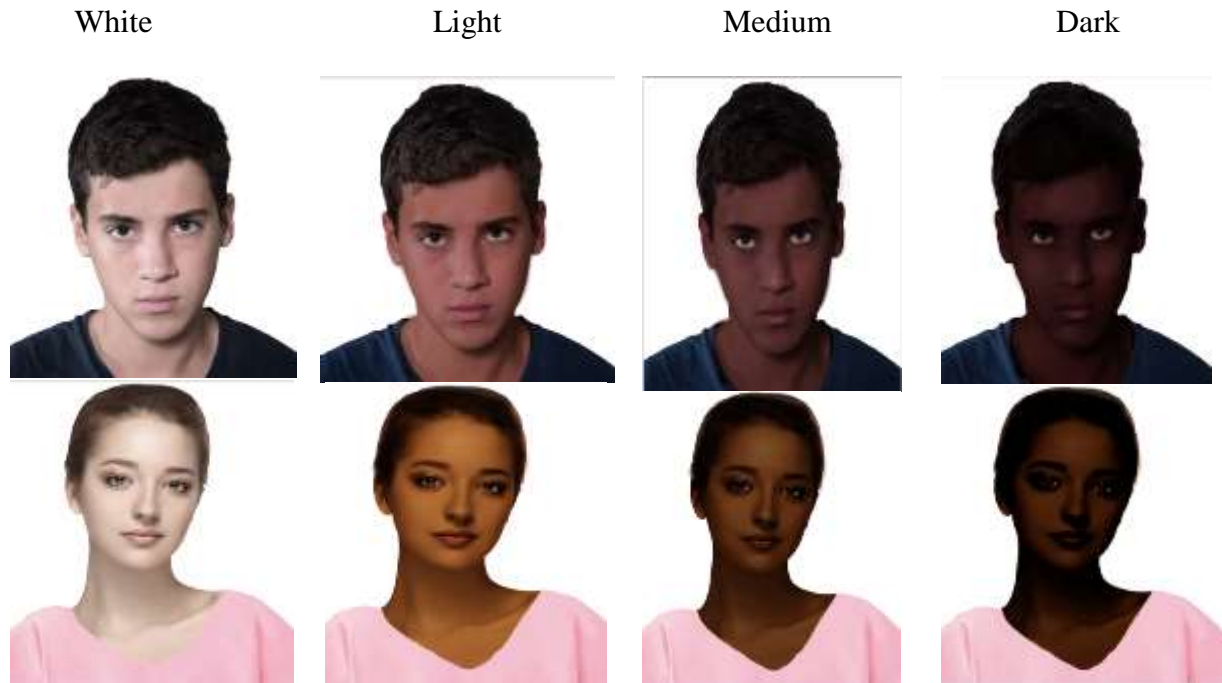


Figure 3.2. Final Student Picture Matrix

Selected images were free to use, meaning the image did not require copyright permission. Skin tones were created within the ability of photoshop software but also using the Massey-Martin Scale as a point of reference. The Massey-Martin Scale (see Figure 3.3) is widely used to measure skin tone and was initially created to acknowledge the broad category of skin tones and the differences in appearance that could impact social outcomes.

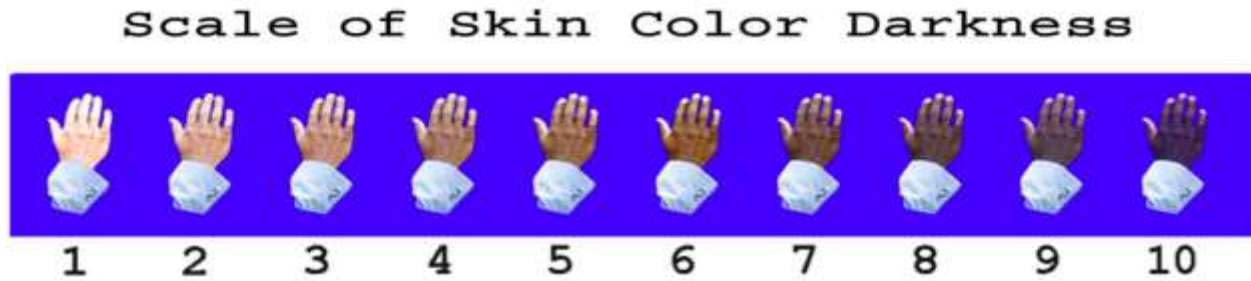


Figure 3.3. Massey-Martin Scale

Following the creation of the new student pictures, six different respondents (educators) were asked to rate the new pictures created in Photoshop. Similar to the feedback given about the first set of student pictures, respondents noted that it was difficult to hypothesize a specific race when viewing the pictures in absence of other phenotypic features, but that the students were white (white students), Hispanic (light-skin black students), Hispanic or black (medium-skin black student) and then finally black (dark-skin black student) for the darkest hue. One educator noted that the pictures address the main purpose of the study, to assess skin tone and not specifically race of student although at times skin tone is linked to race and is difficult to measure separately. Because this is a study of skintone, and not race, and because respondents clearly made distinctions by gradation of color, I used the pictures in Figure 3.2 for the final study.

Dependent Variable

The dependent variable in this dissertation study is the discipline decision assigned by participants to each scenario. A pilot study using the original student pictures was completed to determine educators' discipline decisions based on student skin tone and gender, using three discipline scenarios. The purpose of the pilot study was primarily to determine the appropriateness of the outcome measure and procedure.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Using a convenience sample, twenty-one educators completed one of the nine randomly assigned surveys. Participants were emailed a link to begin the survey and were told that they were participating in a study about classroom behavior management. Participants needed to agree to be a participant and confirm K-12 educator status to continue with the survey. Participants were told that they could discontinue the survey at any time if they did not feel comfortable. Participants were randomly assigned to take one of nine surveys. One version of the survey, not featuring a student picture, asked participants to select an infraction for three different discipline scenarios. The other eight versions of the survey featured the same three discipline scenarios and included one of the eight student pictures. The decision to feature the same student picture throughout each survey was to decrease the likelihood that participants would detect the skin tone condition, which could impact discipline decisions so that the respondent would appear to have given a “socially desirable” response. If only exposed to one student picture, a participant would not know that there are other treatment conditions. A picture of a white, light, medium dark-skinned male or female student was featured on the left side of the survey.

Participants were able to read each scenario once and could not return to previously answered questions, nor could questions be skipped. After reading a scenario, participants determined how likely they were to suspend, verbally reprimand, ignore, or call the student's parent. After reading the three scenarios, participants answered demographic information about themselves (gender, race, school demographics, and years of teaching experience). Finally, participants were given the opportunity to debrief by commenting and providing feedback about the survey. On average, participants took four minutes to complete the survey. Although results from the pilot study were not statistically significant, information gained from the pilot study

was used to revise the outcome variable, discipline decision.

Independent Variable

The independent variable in this study are participants' race and gender. An educator's race and gender could impact discipline decisions. A research question was developed to determine if participant race and gender played a role in the likelihood of assigning a particular infraction based. Determining if there is a pattern in the way educators of various race/gender combinations respond to discipline scenarios can be used in preparation of future professional development and to make educators more aware of race/gender-based response to discipline. Participants will be asked to identify their race and gender at the end of the study, which will then be analyzed to determine if differences exist between the moderating and dependent variable. A description of all variables being considered in this study are outlined in Table 3.1.

Table 3.1

Descriptions of Variables in the Investigation

Name of Variable	Type of Variable	Categories
Student Skin Tone	Categorical	White, Light, Medium, Dark.
Student gender	Categorical	Male or Female
Educator gender	Categorical	Male or Female
Educator race	Categorical	American Indian, Asian, Black, Hispanic, White or Other
Disciplinary decision	Interval	5-point Likert Scale (Very Unlikely, Unlikely, Undecided, Likely and Very Likely)

Post Pilot Study Changes to the Survey

As a result of the pilot study, an additional "dummy" or practice question was added to the beginning of the survey to give participants an opportunity to understand the flow/layout of the study and allows an opportunity to "practice." In the pilot study, the likelihood of an educator

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

selecting “very likely” to suspend a student decreased with each scenario. Placing a scenario at the beginning of the survey that will not be included in the analysis eliminates the chance that participants are selecting “very likely” to suspend because they believe they do not have other discipline options. Another change will be that the survey will be timed and participants will now have fifteen minutes to complete the entire survey. Results from the pilot study showed on average participants completed the pilot study surveys in 4 minutes. Eleven additional minutes were added to the average time to account for the additional questions/lengthier scenarios and to ensure participants are given sufficient time to complete the survey. This timing decision was made to mimic, as closely as possible, the time constraints faced by educators when they decide to make an office referral or to discipline a student. Timed surveys reduce drop-out rates, because participants cannot leave their computers and finish it later once they have started. Timed surveys also capture the first thing that comes to a person’s mind when they are completing a survey. An additional question was added to the final protocol, asking participants if they are K-12 educators, and if so, to identify the title description of their jobs. At the completion of the survey, participants were given the direction to email the researcher with the words “Completed the survey” in the subject line to be entered into the drawing to win a fifty dollar Starbucks gift card.

Data Collection

Respondents were presented with an embedded link to the QuestionPro invitation email. Data was automatically recorded on the website. Participants received a subsequent email, reminding participants to complete the survey. A third reminder email was sent one week before the survey closing. At the end of the survey completion deadline, data was downloaded into an excel spreadsheet to prepare for data entry into SPSS.

Data Coding Procedures

For each scenario, respondents rated how likely they were to select each of the response options (ignore, reprimand, call home, office referral, or suspend) using a 5-point Likert Scale. For each scenario, I selected the harshest discipline decision each participant assigned out of the four questions asked. For example, if a participant stated that she/he was "unlikely" to ignore, reprimand, call home or suspend a student but "very likely" to utilize an office referral, the office referral was coded into a variable to be measured for the most likely/most punitive infraction assigned. Similarly, if a participant stated that he/she was "unlikely" to ignore, reprimand or call home, but "very likely" to make an office referral and "very likely" to recommend a student for suspension, the harshest infraction (recommend for suspension) was coded for the new variable.

Participant responses were recoded into a scale where 1 = least punitive or less harsh to 5 = most punitive or harshest. Although scenario one was created as a practice question, I analyzed responses to scenario one to determine if a statistically significant difference existed. Ultimately, results were analyzed and were coded from least to most punitive, rather than the specific response. One of the limitations of this study, as discussed in chapter 3, is response rate as it relates to cell size. The surveys were distributed to 1,042 participants during the first wave and 1,833 in the second wave. Out of 2,875 participants, only 228 responded to the survey. Due to the

number of treatments surveys (9), some of the cell sizes were too small to make meaningful conclusions. For instance, because the male teacher overall response rate was 40, the number of male responses for each of the treatments are too small to provide information on the interaction of respondent gender and disciplinary decisions by student skin tone. The same is true for respondent race. I could not control how many people responded to a survey when all surveys were sent to approximately the same number of people, and I expanded and oversampled for males and teachers of color. Thus, my final response rate, although adequate for the overall analysis, does not allow for an exploration of whether the race and/or gender of the respondent was a factor in the discipline decision. Furthermore, response rates for each scenario declined slightly due to survey dropout rates.

Data Analysis

SPSS 26.0 was used to analyze data. Means, standard deviations, frequencies and other descriptive statistics were determined for each variable in this study. Each research question was analyzed using the appropriate statistical test. Table 3.2 contains a description of the selected statistical tests that were used to analyze data collected in this study.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Table 3.2

Research Question Matrix

Research Question	Statistical Test
How "harshly" do educators discipline based on scenario type?	Descriptive
Are there difference in educators discipline decisions by color of vignette student skin tone?	ANOVA
Are there differences in educator discipline decision by gender of student vignette?	ANOVA
Are there differences in educator discipline decisions by intersection of gender and skin tone of vignette student?	ANOVA
Are there differences in educators discipline decision by race of educators?	ANOVA
Are there differences in educator discipline decision by race of educator and skin tone of vignette student?	ANOVA
Are there differences in discipline decisions made by black and white educators based upon brown skin tones of light, medium, dark?	ANOVA
Are there differences in educator discipline decisions by race of educator and gender of vignette student?	ANOVA
Are there differences in educator discipline decision by gender of educator?	ANOVA
Are there differences in educator discipline decisions by gender of the educator and vignette student skin tone?	ANOVA
Are there differences in educator discipline decision by gender of educator gender of vignette student?	ANOVA

School Discipline Scenarios

Practice question. You are preparing your class to take a major state-wide test. Over the last few weeks, you have constantly reminded Student A not to interrupt others while they are taking the test. Student A is often impatient and constantly calls out in class. This behavior is annoying to you and to many of the other students in your class. You frequently remind Student A that these behaviors are frustrating, but the calling out and interrupting persists.

Defiance of authority/insubordination. Over the last few weeks, you have noticed a change in Student A behaviors including constantly lying to you and not being able to complete an entire assignment during the class period. You have sent multiple letters home to Student A's parents about incomplete homework assignments and coming to class without the textbook or notebook. You have also directly addressed excessive lying with Student A on multiple occasions, in which student A responds, "It won't happen again."

Today during class, Student A asked to go to the clinic because of suspicion of a concussion in gym class that was not reported to the gym teacher. You ask Student B to follow Student A to the clinic to ensure safe arrival. Student B reported that Student A went into the restroom, then into the teacher's workroom, and then ran off and could not be found. When Student A returned to the class. Student A reported going to the clinic. But when questioned, admitted to lying to get out of class.

Classroom or campus disruption. You are teaching a lesson that students will later be tested on at the end of the quarter. Student A is playing with various items under the desk and talking non-stop during your lecture. This behavior is common for Student A. You have spoken to Student A several times, trying to get the student to pay attention in class or at least not disrupt other students.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

You tell student A to put the items in the trash and to pay attention to your presentation. Student A walks to the trash and throws the first item away, then returns and gets the second item from the desk. Student A then walks that item to the trash and throws it away. This continues two more times. Once seated, student A continues to talk to other students, asking questions and making comments. Classmates appear to be trying to ignore Student A.

Student A continues and as other students start to ask Student A to be quiet, Student A calls one student a “Bitch” and responds to another student who asked for the behavior to stop with “Fuck you.” Throughout this behavior, you have told Student A to sit down and be quiet three times. Finally, Student A sits down but then begins popping bubble gum. Each time, Student A blows the bubble and pops it even louder.

Disruptive demonstrations. It is Monday morning, and the class is starting a new chapter in the textbook. Student A’s head is on the desk, and Student A is sleeping. This is not unusual for Student A. Neither is interrupting class. Both things have happened regularly in class, and you have talked with Student A several times about the behavior.

You tell student A to “sit up and pay attention.” Student A responds to your request by sitting up, but begins to make loud, random sounds. Some classmates tell Student A to “cut it out”. Student A immediately discontinues the noise. However, within minutes, Student A begins to make inappropriate jokes while you are giving directions for homework. Student A then begins again, making loud noises. You turn to Student A and say, “Stop now, or you will need to go to the principal’s office.” Student A replies by mocking you, saying “Stop now or you will have to go to the principal’s office.”

VCU IRB

Prior to collecting data, this dissertation study was approved by the Virginia Commonwealth University Institutional Review Board (IRB), consistent with the rules and regulations of the institution.

Limitations

The participants in this study work in various school divisions in the southeastern region of the United States. Therefore, results cannot be generalized to other settings or states. For the purpose of this study, an educator is defined as anyone who has or currently works in the k-12 setting and can make a student discipline referral. However, the criteria of who can make a discipline referral may vary between school divisions. This study does not analyze the district level policies about making office referrals. Additionally, participants work at different types of public schools. For example, some teachers may work at Title I or specialty schools whereas others may work at more affluent or National Blue Ribbon Schools. Although participants are asked to identify the majority of their student population (socioeconomic status), this question does not account for other characteristics that make each school unique.

Additionally, a person's skin tone is subjective and one's degree of lightness or darkness on an individual's perception. The skin tones depicted in the students' pictures in this study may not be perceived the same by all participants. For example, African Americans or people with more exposure to people of color may view skin tone differently than some who is white or does not have much experience viewing people of color. Participants utilized technology to complete this survey and different devices can cause students' skin tones to appear darker or lighter than intended.

The Massey-Martin Scale is criticized for having low inter-coder reliability.

Using data from the 2012-2014 General Social Survey, Hannon & DeFina (2015) found that when analyzing ratings of the same participants and interviewers, African American respondents are rated as having a higher than average skin darkness when rated by white interviewers than when rated by African American interviewers. Less than a quarter of African American respondents listed in the top three categories of skin darkness in 2012 were listed in the top three categories in 2014, meaning that skin tone is subjective and is often inaccurately measured or categorized even when using the same participants and interviewers. Although this skin tone measurement tool has flaws, no other scale of its type exists. Therefore, I am using this measure as a point of reference for the creation of student skin tones and not as a definitive guideline.

Delimitations

Since the study was not conducted in the real educational environment, the complexities that exist within classrooms cannot be captured. Some factors may occur in everyday classrooms that cannot be replicated in this study. Elementary and middle school students have lower discipline rates compared to high school students. High school students are the only group mentioned in the discipline scenarios because high schoolers are most impacted by school disciplining issues (Brookings Institute, 2017). Additionally, although colorism exists in other racial/ethnic groups, this study highlights colorism as it pertains to African American students because they are the most frequently disciplined group.

Assumptions

In this proposed study, the following assumptions have been made: All educators:

1. Have experience with disciplining students and with classroom management.
2. Do not purposely or knowingly discriminate/demonstrate bias against specific students
3. Believe that all students should be treated equally and fairly.

Researcher's Perspective

As an African American female, I have witnessed and experienced colorism first hand. I believe due to my life experiences, that colorism is a subtle form of discrimination that is difficult to measure. I also think that all people demonstrate bias in one area or another. However, I believe that educators unknowingly display implicit bias towards students. Most educators enter the field to make a positive impact in the lives of their students. I admit that as an educator I hold various forms of bias that I did not have to confront before initiating this work. As a result of the recognition of my own personal biases, I have joined my school division's equity committee to "unpack" the biases that I hold that could impact the educational and social-emotional outcomes of the students that I serve in the school setting. Meeting frequently throughout the year, the committee consists of teachers, principals, and community leaders. We discuss and research useful tools educators can use to develop cultural competency and to confront equity issues in the school division. My research is a contribution, not just to academia, but to the larger society. This study will bring awareness to the idea that colorism in the public-school setting may occur during the process of disciplining students. This study is an effort to start a critical discussion about this vital issue and to help school divisions provide effective professional development to educators that address colorism as a form of discrimination and

implicit bias.

Disclosures

The author of this proposed dissertation is currently an employee at Chesterfield County Public Schools. This study is not funded by a grant. The Starbucks gift card will be purchased with the researcher's personal funds to increase participant's likelihood of completing the survey.

CHAPTER 4: FINDINGS

My study simulates real-life classroom interactions between students and educators to examine how educators report they would discipline students. Students in the simulated vignette represent one of four skin tones. Previous research has shown that black students are suspended at a disproportionately higher rate than same-aged peers (Skiba, et. al., 2011). Much of the research continues to focus on the interracial gap with disregard for the intra-racial gap. The research questions outlined in this chapter focus on intra-racial patterns to better understand how race and gender of the student and educator intersect with students' skin tone.

As described in Chapter 3, there are four scenarios (including a practice scenario that I have analyzed). The scenarios were created with the feedback from educators based upon current discipline data from the Department of Education in a state in the southeast United States. In no particular order, the top three infractions identified by the state department of education were:

1. Classroom/Campus Disruption
2. Defiance of Authority/Insubordination
3. Disruptive Demonstrations

In schools in this state, when there is a discipline report, school administrators are responsible for coding each infraction out of more than thirty options. Among these options, several are more straightforward than others; for instance, offenses involving fighting, possession of firearms, or possession of drugs. However, for many infractions, the coding of the the discipline decision is more subjective. One administrator may code an infraction as disruptive demonstration, while another administrator would code the same infraction as defiance of authority. Both ways of coding may be accurate. This study examined three of the more

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

subjective categories. The scenarios for this study focus on three categories: classroom/campus disruption (scenario one/practice and scenario four),; defiance of authority (scenario two),; and disruptive demonstration (scenario three).

These scenarios were delivered in an online survey to which 228 K-12 educators responded to an online survey. Over half of the participants stated that they worked at a school where more than 50% of students qualified for free and reduce lunch. For all but the control (no picture), each respondent recorded a disciplinary action for a student who was represented one of eight conditions (male or female x white skin tone, light brown skin tone medium brown skin tone, and dark brown skin tone). Participant demographic data are presented in Table 4.1 and Table 4.2. Tables 4.1 and 4.2 show a total of 202 instead of 228 participants because 26 respondents did not complete the demographic questions, which were presented at the end of the study.

Table 4.1

Participant Gender

Male		Female		Missing		Total
n	%	n	%	n	%	
40	17.5	162	71.1	26	11.4	228

Table 4.2

Participant Race

	n	%
Hispanic or Latino/a	5	2.2
American Indian	2	.9
Asian	4	1.8
African American or Black	39	17.1
White	143	62.7
Other	5	2.2
More than one	4	1.8

Missing	26	11.3
Total	202	100

Research Question 1

How "harshly" do educators discipline based on scenario type?

Prior to examining differences by skin tone and other variables, I examined the disciplinary responses for each scenario. Classroom/campus disruptions (scenarios one and four) received the least punitive responses with 3.5% of respondents indicated suspension for scenario 1 and 7.9% indicating suspension for scenario 4. Both of these classroom/campus disruption scenarios had the highest percentage of teachers who said they would ignore the student behavior, compared to 10.1% of respondents recommending a suspension for defiance of authority (scenario two) and 17.5% recommending suspension for disruptive demonstration (scenario three). The defiance of authority vignette, scenario two, 10.1% of educators selected the most punitive decision. Scenario three yielded the highest percentage of educators selecting the most punitive decision at 17.5%. Overall, most educators decided to call the student's parents (scenarios one and two) or make an office referral (scenarios three and four) Table 4.3 details the frequency of each discipline decision for each scenario.

Table 4.3

Frequency of Discipline

	Ignore	%	Reprimand	%	Call	%	Office Referral	%	Suspension	%
Scenario 1	8	3.5	47	20.6	137	60.1	28	12.3	8	3.5
Scenario 2	1	.4	10	4.4	95	41.7	87	38.2	23	10.1
Scenario 3	0	0	11	4.8	48	21.1	113	49.6	40	17.5
Scenario 4	13	5.7	22	9.6	69	30.3	84	7.9	18	7.9

Research Question 2

Are there differences in educator discipline decisions by vignette student skin tone?

A. Comparison of all skin tones and control group. First, I wanted to understand how "harshly" educators disciplined students across the four skin tone groups (light, medium, dark, white) in addition to the control group, a total of five categories (no gender distinctions, male and female students combined). In three of the four scenarios, there were no statistically significant differences by skin tone. There was no statistically significant difference for scenario one, $F(4,223) = 1.003, p = .41$; scenario three, $F(4,207) = 2.106, p = .08$; or scenario four, $F(4,201) = 1.770, p = .14$. There was a statistically significant difference in the discipline decisions assigned to students in scenario two, $F(4,211) = 2.579, p = .04, \eta^2 = .05$. However post-hoc analysis did not show a statistically significant difference between any of the skin tone/control means.

B. Comparison of all skin tones without a control group. Next, I wanted to understand how "harshly" educators disciplined students across the four skin tone groups (light, medium, dark, white) without the control group. There was no statistically significant difference for scenario one, $F(3,185) = 1.233, p = .30$; scenario two, $F(3,174) = 1.733, p = .16$; scenario three, $F(3,170) = 1.998, p = .12$; or scenario four, $F(3,167) = 2.445, p = .07$.

C. Comparison of light, medium, and dark students with the control group. In two of the four scenarios, there were no statistically significant differences by skin tone. There was no statistically significant difference for scenario one, $F(3,187) = 1.178, p = .32$, or scenario four, $F(3,168) = .998, p = .40$. There was a statistically significant difference in the discipline decisions assigned to students in scenario two, $F(3,178) = 3.022, p = .03, \eta^2 = .05$, and scenario three, $F(3,174) = 2.784, p = .04, \eta^2 = .05$. In scenario two, post-hoc analysis revealed light

skin students ($M = 3.73$, $SD = .811$) were disciplined more "harshly" than the control group ($M = 3.31$, $SD = .661$). However, the variance accounted for (.05) is not meaningful. Post-hoc analysis did not show a statistically significant difference between any category means for scenario three.

D. Comparison of light, medium, and dark students without the control group. I then decided to complete an analysis omitting white student vignettes as well as the control group to determine how "harshly" educators disciplined students across the three brown skin tone groups (light, medium, dark). There was no statistically significant difference for scenario one, $F(2,149) = 1.540$, $p = .22$; scenario two, $F(2,141) = 1.212$, $p = .30$; scenario three, $F(2,137) = 2.611$, $p = .08$; or scenario four, $F(2,134) = 1.229$, $p = .30$.

E. Comparison of white students with the control group. Finally, to answer this research question in all aspects, I decided to complete an analysis to determine how "harshly" educators disciplined students, including only white student vignettes versus the control group. There was no statistically significant difference for scenario one, $F(1,74) = .041$, $p = .84$; scenario two, $F(1,70) = .301$, $p = .59$; scenario three, $F(1,70) = .103$, $p = .75$; or scenario four, $F(1,67) = .902$, $p = .35$. Sample sizes, means and standard deviations for student all skin tones and control group based on "harshness" of discipline decisions for each scenario are presented in Table 4.4.

Table 4.4

Sample Size, Mean and Standard Deviation for Student Skin Tone and Discipline Decision

Skin Tone	Scenario One			Scenario Two			Scenario Three			Scenario Four		
	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
Light	52	2.96	.815	49	3.73	.811	48	4.08	.767	47	3.30	.976
Medium	50	3.08	.804	48	3.52	.772	47	3.96	.806	47	3.60	.948
Dark	50	2.80	.782	47	3.72	.682	45	3.71	.815	43	3.44	.825
White	37	2.84	.764	34	3.41	.821	34	3.76	.890	34	3.03	1.114
No Picture	39	2.97	.695	38	3.32	.662	38	3.71	.515	35	3.29	1.126

F. Comparison of Students of Color, White Students and Control Group. After analyzing student skin tones in every possible way, I decided to create a new student of color variable (including light, medium and dark skin tones). I analyzed the student of color category, white student category and the control group category. There was no statistically significant difference for scenario one, $F(2,225) = .371, p = .69$; scenario three, $F(2,209) = 1.397, p = .25$; or scenario four, $F(2,203) = 2.478, p = .09$. Results showed a statistically significant difference for scenario two, $F(2,213) = 3.920, p = .02, \eta^2 = .04$. A Bonferroni post-hoc criterion for scenario two showed educators disciplined students of color ($M = 3.66, SD = .759$) more "harshly" than the control group ($M = 3.32, SD = .662$). However, the variance accounted for (.04) was not meaningful. Analysis of sample size, means, and standard deviation of each scenario is presented in Table 4.5.

Table 4.5

Sample Size, Mean and Standard Deviation for Student Skin Tone, Control Group and Discipline Decision

Skin Tone	Scenario One			Scenario Two			Scenario Three			Scenario Four		
	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
Student of Color	152	2.96	.804	144	3.66	.759	140	3.92	.805	137	3.45	.923
White	37	2.84	.764	34	3.41	.821	34	3.76	.890	34	3.02	1.11
Control Group	39	2.87	.705	38	3.32	.662	38	3.71	.515	35	3.29	1.13

G. Comparison of students of color and control group. Next, I decided to complete another analysis, omitting the white student variable and only considering students of color and the control group in the analysis. There was no statistically significant difference for scenario one, $F(1,189) = .289, p = .592$; scenario three, $F(1,176) = 2.34, p = .13$; or scenario four, $F(1,170) = .759, p = .39$. Results showed a statistically significant difference for scenario two, $F(1,180) = 6.50, p = .01, \eta^2 = .03$. Comparing means for scenario two, students of color ($M = 3.66, SD = .759$) were disciplined more "harshly" than the control group ($M = 3.32, SD = .821$). However, the variance accounted for (.03) was not meaningful. Analysis of sample size, means and standard deviation for each scenario is presented in Table 4.6.

Table 4.6

Sample Size, Mean and Standard Deviation for Student of Color, Control Group and Discipline Decision

Skin Tone	Scenario One			Scenario Two			Scenario Three			Scenario Four		
	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
Students of Color	152	2.95	.804	144	3.66	.759	140	3.92	.805	137	3.45	.923
Control	39	2.87	.696	38	3.32	.821	38	3.71	.515	35	3.29	1.13

H. Comparison of students of color and white students. I then decided to complete another analysis, omitting the control group variable and only considering students of color and white students in the analysis. There was no statistically significant difference for scenario one, $F(1,187) = .563, p = .45$; scenario two, $F(1,176) = 2.847, p = .09$; or scenario three, $F(1,172) = .994, p = .32$. Results showed a statistically significant difference for scenario four, $F(1,169) = 5.08, p = .03, \eta^2 = .03$. Comparing means for scenario four, students of color ($M = 3.45, SD = .923$) were disciplined more "harshly" than white students ($M = 3.03, SD = 1.114$). However, the variance accounted for (.03) was not meaningful. Analysis of sample size, means, and standard

deviation, or each scenario is presented in Table 4.7.

Table 4.7

Sample Size, Mean and Standard Deviation for Student of Color and White Students

Skin Tone	Scenario One			Scenario Two			Scenario Three			Scenario Four		
	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
Students of Color	152	2.95	.804	144	3.66	.759	140	3.92	.805	137	3.45	.923
White	37	2.84	.764	34	3.41	.821	34	3.76	.890	34	3.02	1.114

I. Comparison of all students and control group. Finally, I created a new variable by combining all student skin color categories to determine if there was a difference between discipline decisions for all students, regardless of skin tone, when compared to the control group. There was no statistically significant difference for scenario one, $F(1,226) = .156, p = .69$; scenario three, $F(1,210) = 1.680, p = .196$; or scenario four, $F(1,204) = .171, p = .68$. A statistically significant difference existed in scenario two, $F(1,214) = 4.813, p = .03, \eta^2 = .02$. Comparing means for scenario two, all students ($M = 3.61, SD = .775$) were disciplined more "harshly" than the control group ($M = 3.32, SD = .662$). However, the variance accounted for (.02) was not meaningful. Analysis of sample size, means and standard deviation or each scenario is presented in Table 4.8.

Table 4.8

Sample Size, Mean and Standard Deviation for All Students and Control Group

Skin Tone	Scenario One			Scenario Two			Scenario Three			Scenario Four		
	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
All Students	189	2.93	.795	178	3.61	.775	174	3.89	.822	171	3.36	.975
Control Group	39	2.87	.764	38	3.32	.662	38	3.71	.515	35	3.29	1.13

To increase understanding of results and to provide a quick reference for readers, Table 4.9 outlines the statistically significant results for question two, parts A through H.

Table 4.9

Summary of Statistically Significant Differences for Question Two Part A through Part H

Scenario 1: No statistically significant differences
Scenario 2: Statistically significant differences in questions A,C,F, G and I, not meaningful
Scenario 3: No statistically significant differences
Scenario 4: Statistically significant differences in questions H, not meaningful

Research Question 3

Are there differences in educator discipline decisions by gender of vignette student?

An analysis was completed to determine if educators disciplined students more "harshly" based on gender. Results showed that there were no statistically significant differences in how "harshly" educators assigned consequences based on student gender for scenario one, $F(1,187)=.275, p =.60$; scenario two, $F(1,176)=1.142, p =.287$; scenario three, $F(1,172)=.251, p =.62$; or scenario four, $F(1,169)=1.385, p =.24$. Sample sizes, means and standard deviations for student gender based on most likely discipline decisions for each scenario is presented in Table 4.10.

Table 4.10

Sample Size, Mean and Standard Deviation for Student Gender and Discipline Decision

	<u>Scenario One</u>			<u>Scenario Two</u>			<u>Scenario Three</u>			<u>Scenario Four</u>		
Student Gender	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
Male	83	2.89	.733	76	3.68	.716	76	3.86	.743	76	3.46	.986
Female	106	2.95	.843	102	3.56	.815	98	3.91	.881	95	3.29	.964

To increase understanding of results and to provide a quick reference for readers, Table 4.11 outlines the statistically significant results for question three.

Table 4.11

Summary of Statistically Significant Differences for Question Three

Scenario 1: No statistically significant differences
Scenario 2: No statistically significant differences
Scenario 3: No statistically significant differences
Scenario 4: No statistically significant differences

Research Question 4

Are there differences in educator discipline decisions by intersection of gender and skin tone of vignette student?

A. Comparison of discipline decisions for students of color, white students

and control group by intersection of gender. To determine how "harshly" educators disciplined students when considering treatment type (gender x skin tone), I completed an analysis including students of color, white students, and the control group. There was no statistically significant difference between the variables. I then decided to remove the control group category to see if there were differences in the ways student disciplinary decisions were made. There were no statistically significant differences in scenario one, $F(7,163)=2.464, p = .02$; scenario two, $F(7,163)=2.464, p = .02$; or scenario three, $F(7,163)=2.464, p = .02$. There was a statistically significant difference in harshness of discipline decisions for scenario four, $F(7,163)=2.464, p = .02, \eta^2 = .09$. Medium skin female students ($M = 4.00, SD = .667$) were disciplined more "harshly" than white female students ($M = 2.82, SD = .568$) in scenario 4.

B. Comparison of discipline decisions for students of color by gender. I then compared only decisions about students of color, without the control group and without white

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

students. There was not a statistically significant difference in the discipline decisions assigned to students in scenario two, $F(5,138)=.781, p=.55$. There was a statistically significant differences in scenario one, $F(5,146)=2.586, p=.03$ and scenario four, $F(5,131)=2.288, p=.05$. However post-hoc analysis for scenario one and four did not show a statistically significant difference between the category means. A statistically significant difference was found in scenario three, $F(5,134)=2.435, p=.04, \eta^2=.08$. A Bonferroni post-hoc criterion for significance was completed to further understand differences in how "harshly" students were disciplined by educators for scenario three. In scenario three, educators disciplined light skin males ($M=4.30, SD=.775$) more "harshly" than dark skin males ($M=3.58, SD=.857$). Analysis of sample size, means and standard deviation, or each scenario is presented in Table 4.12.

Table 4.12

Sample Size, Mean and Standard Deviation for Student Skin Tone/Gender and "Harshness" of Discipline

Skin Tone	Scenario One			Scenario Two			Scenario Three			Scenario Four		
	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
White Female	19	2.89	.875	17	3.53	.662	17	3.65	.862	17	2.82	.568
Light Female	22	2.68	.568	21	3.76	.700	21	3.81	.680	21	3.29	.695
Medium Female	20	3.35	.587	19	3.68	.820	19	4.05	.705	19	4.00	.667
Dark Female	22	2.68	.716	19	3.74	.562	19	3.89	.737	19	3.68	.875
White Male	18	2.78	.647	17	3.29	.849	17	3.88	.928	17	3.24	.970
Light Male	30	3.17	.913	28	3.71	.897	27	4.30	.775	26	3.31	1.050
Medium Male	30	2.90	.885	29	3.41	.733	28	3.89	.165	28	3.32	1.020
Dark Male	28	2.89	.832	28	3.71	.763	26	3.58	.857	24	3.25	.84
No Picture	39	2.87	.695	38	3.32	.662	38	3.71	.515	35	3.29	.716

To increase understanding of results and to provide a quick reference for readers, Table 4.13 outlines the statistically significant results for question four A and B.

Table 4.13

Summary of Statistically Significant Differences for Question Four Part A and Part B

Scenario 1: No statistically significant differences
Scenario 2: No statistically significant differences
Scenario 3: Statistically significant, meaningful
Scenario 4: Statistically significant, meaningful

Question 5

Are there differences in educator discipline decisions by race of educator and skin tone of vignette student?

This research question asked if an educator's race is related to discipline decisions for students based on the intersection of student vignette skin tone and gender. Due to small sample sizes, I could not answer this question. Therefore, I replaced it with a question that I could explore:

Question 6

Are there differences in educator discipline decisions by race of educator?

Analysis determined that there were no statistically significant differences between "harshness" of educator decisions based on the race of the educator in scenario one, $F(1,180)=.338, p =.56$; scenario two, $F(1,180)=.01, p =.95$; scenario three, $F(1,180)=1.57, p =.21$; or scenario four, $F(1,180)=.127, p =.72$. Sample size, means, and standard deviation for each scenario are presented in Tables 4.14.

Table 4.14

Sample, Mean and Standard Deviation by Race of Educator and Discipline Decision

	<u>Participant Race</u>					
	Black			White		
	n	Mean	SD	n	Mean	SD
Scenario 1	39	2.97	1.01	143	2.90	.668
Scenario 2	39	3.59	.785	143	3.58	.745
Scenario 3	39	3.95	.826	143	3.78	.745
Scenario 4	39	3.30	.950	143	3.37	.984

Question 7

Are there differences in the discipline decisions made by black and white educators based upon brown skin tones of light, medium, and dark of student vignette?

Analysis determined that there were no statistically significant differences between "harshness" of educator decisions by race of educator and student skin tone in scenario one, $F(2,142)=.496, p=.69$; scenario two, $F(2,142)=.57, p=.63$; scenario three, $F(2,142)=.638, p=.59$; or scenario four, $F(2,142)=1.262, p=.29$. Sample size, means, and standard deviation of each scenario are presented in Tables 4.15-4.19.

Table 4.15

Sample, Mean and Standard Deviation by Race of Educator and Student Skin Tone for Scenario 1

	<u>Participant Race</u>					
	Black			White		
Student Skin Tone	n	Mean	SD	n	Mean	SD
Light	10	3.00	.943	29	2.97	.865
Medium	13	3.15	.899	28	3.00	.720
Dark	7	2.57	.787	32	2.91	.588

Table 4.16

Sample, Mean and Standard Deviation by Race of Educator and Student Skin Tone for Scenario 2

	<u>Participant Race</u>					
	Black			White		
Student Skin Tone	n	Mean	SD	n	Mean	SD
Light	10	4.00	.666	29	3.79	.726
Medium	13	3.62	.870	28	3.54	.744
Dark	7	3.86	.690	32	3.69	.644

Table 4.17

Sample, Mean and Standard Deviation by Race of Educator and Student Skin Tone for Scenario 3

	<u>Participant Race</u>					
	Black			White		
Student Skin Tone	n	Mean	SD	n	Mean	SD
Light	10	4.30	.823	29	3.97	.731
Medium	13	4.00	.707	28	3.89	.832
Dark	7	4.00	1.000	32	3.66	.745

Table 4.18

Sample, Mean and Standard Deviation by Race of Educator and Student Skin Tone for Scenario 4

	<u>Participant Race</u>					
	Black			White		
Student Skin Tone	n	Mean	SD	n	Mean	SD
Light	10	2.80	.919	29	3.45	.985
Medium	13	3.85	.801	28	3.61	.956
Dark	7	3.29	.951	32	3.47	.621

Table 4.19

Summary of Statistically Significant Differences for Question Seven

Scenario 1: No statistically significant differences
Scenario 2: No statistically significant differences
Scenario 3: No statistically significant differences
Scenario 4: No statistically significant differences

Question 8

Are there differences in educator discipline decisions by race of educator and gender of vignette student?

Analysis showed a statistically significant difference between educator race and student gender in scenario one, $F(1,146)=5.215, p=.024, \eta^2=.03$. Comparison of the means (Table 20) show that black educators discipline female students more "harshly" ($M=3.21, SD=.787$) than white educators ($M=2.88, SD=.558$). However, the variance accounted (.03) for was not meaningful. There were no statistically significant differences by race of educator in the ratings of male students.

Table 4.20

Sample, Mean and Standard Deviation by Race of Educator and Student Gender Scenario 1

	<u>Educator's Race</u>					
	Black			White		
Student Gender	n	Mean	SD	n	Mean	SD
Male	14	2.57	1.016	67	2.93	.805
Female	19	3.21	.787	50	2.88	.558

There were no statistically significant differences in discipline decision based on educator's gender and student gender for scenario two, $F(1,146)=.077, p=.782$; or scenario three, $F(1,146)=1.244, p=.267$. Sample, mean and standard deviation data for scenarios two and three are provided in Tables 4.21 and 4.22.

Table 4.21

Sample, Mean and Standard Deviation by Race of Educator and Student Gender for Scenario 2

	<u>Educator's Race</u>					
	Black			White		
Student Gender	n	Mean	SD	n	Mean	SD
Male	14	3.64	.633	33	3.60	.799
Female	19	3.79	.855	50	3.66	.688

Table 4.22

Sample, Mean and Standard Deviation by Race of Educator and Student Gender for Scenario 3

	<u>Educator's Race</u>					
	Black			White		
Student Gender	n	Mean	SD	n	Mean	SD
Male	14	3.93	1.072	67	3.88	.826
Female	19	4.11	.658	50	3.70	.735

For scenario four, there was a statistically significant difference by race of participant and gender of student in how "harshly" educators disciplined students, $F(1,146)=8.315$, $p=.005, \eta^2=.05$. Black educators discipline female students ($M=3.74, SD=.872$) more "harshly" than male students ($M=2.79, SD=.802$) (Table 4.23). However, the variance accounted for (.05) was not meaningful.

Table 4.23

Sample, Mean and Standard Deviation by Race of Educator and Student Gender Scenario 4

Student Gender	Educator Race					
	Black			White		
	n	Mean	SD	n	Mean	SD
Male	14	2.79	.802	67	3.44	.942
Female	19	3.74	.872	50	3.34	.939

To increase understanding of results and to provide a quick reference for readers, Table 4.24 outlines the statistically significant results for question eight.

Table 4.24

Summary of Statistically Significant Differences for Question Eight

Scenario 1: Statistically significant, not meaningful
Scenario 2: No statistically significant differences
Scenario 3: No statistically significant differences
Scenario 4: Statistically significant, not meaningful

Question 9

Are there differences in educator discipline decisions by gender of educator?

To understand if an educator's gender made a difference in how students in all vignettes were disciplined, a Univariate Analysis of Variance was used. There was no statistically significant difference for scenario one, $F(1,200)=.695, p=.41$; scenario three, $F(1,200)=.445, p=.506$; or scenario four, $F(1,200)=3.10, p=.08$. Analysis showed a statistically significant difference between educator's gender and how "harshly" students were disciplined for scenario two, $F(1,200)=5.674, p=.02$. Comparison of the means (Table 4.25) show that male educators discipline students more "harshly" ($M = 3.83, SD = .712$) than female educators ($M = 3.52, SD = .733$). However, the variance accounted for (.02) was not meaningful.

Table 4.25

Sample Size, Mean and Standard Deviation for Educator Gender and Discipline Decision

	<u>Scenario One</u>			<u>Scenario Two</u>			<u>Scenario Three</u>			<u>Scenario Four</u>		
Educator Gender	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD
Male	40	3.00	.784	40	3.83	.712	40	3.93	.859	40	3.60	.982
Female	162	2.88	.748	162	3.52	.733	162	3.91	.758	162	3.29	1.01

To increase understanding of results and to provide a quick reference for readers, Table 4.26 outlines the statistically significant results for question nine.

Table 4.26

Summary of Statistically Significant Differences for Question Nine

- Scenario 1: No statistically significant differences
- Scenario 2: Statistically significant, not meaningful
- Scenario 3: No statistically significant differences
- Scenario 4: No statistically significant differences

Question 10

Are there differences in educator discipline decisions by gender of the educator and vignette student skin tone?

Question ten was aimed at understanding if an educator's gender impacts discipline decisions. Due to small cell sizes, it was not possible to analyze an educator's gender and the interaction between student skin tone and gender. Additionally, dark skin student vignette was not removed from the study, although it has a small sample size. Therefore, I examined each of these separately. Analysis showed that for scenario one, $F(3,160)=1.721, p =.17$; scenario three, $F(3,160)=.939, p =.42$; or scenario four, $F(3,160)=.793, p =.50$ there were no statistically significant differences in the discipline decisions for educators regardless of student gender and

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

student skin tone. Sample sizes, means, and standard deviations are listed in Tables 4.27 through 4.29.

Table 4.27

Sample, Mean and Standard Deviation by Gender of Educator and Student Skin Tone for Scenario 1

	Participant Gender					
	Male			Female		
Student Skin Tone	n	Mean	SD	n	Mean	SD
Light	18	2.94	.802	28	2.96	.922
Medium	13	3.46	.660	34	2.91	.753
Dark	4	2.75	.500	39	2.85	.709
White	6	2.50	.837	26	2.80	.694

Table 4.28

Sample, Mean and Standard Deviation by Gender of Educator and Student Skin Tone for Scenario 3

	Participant Gender					
	Male			Female		
Student Skin Tone	n	Mean	SD	n	Mean	SD
Light	18	4.17	.707	28	4.04	.838
Medium	13	4.08	.862	34	3.91	.793
Dark	4	3.25	.500	39	3.79	.801
White	6	3.50	1.225	26	3.77	.815

Table 4.29

Sample, Mean and Standard Deviation by Gender of Educator and Student Skin Tone for Scenario 4

	Participant Gender					
	Male			Female		
Student Skin Tone	n	Mean	SD	n	Mean	SD
Light	18	3.33	.979	28	3.25	1.005
Medium	13	4.077	.760	34	3.41	.957
Dark	4	3.75	.500	39	3.41	.850
White	6	3.00	1.41	26	3.00	1.095

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

However, in scenario two, there was a statistically significant difference in how male and female educators disciplined students based on skin tone, $F(3,160)= 4.247, p =.006, \eta^2 =.07$.

I found that male educators disciplined light-skinned students ($M = 4.06, SD =.63914$) more "harshly" than did female educators ($M = 3.68, SD =.72283$). Sample sizes, means, and standard deviations are listed in Tables 4.30.

Table 4.30

Sample, Mean and Standard Deviation by Gender of Educator and Student Skin Tone for Scenario 2

	Participant Gender					
	Male			Female		
Student Skin Tone	n	Mean	SD	n	Mean	SD
Light	18	4.06	.63914	28	3.68	.772283
Medium	13	3.46	.77625	34	3.59	.74336
Dark	4	3.50	.57735	39	3.77	.66734
White	6	4.33	.51640	26	3.19	.74936

To increase understanding of results and to provide a quick reference for readers, Table 4.31 outlines the statistically significant results for question seven.

Table 4.31

Summary of Statistically Significant Differences for Question Ten

Scenario 1: No statistically significant differences
Scenario 2: Statistically significant, meaningful
Scenario 3: No statistically significant differences
Scenario 4: No statistically significant differences

Question 11

Are there differences in educator discipline decisions by gender of the educator and gender of vignette student?

An analysis was completed to determine if an educator's gender may or may not impact discipline decisions for male and female students. Results showed that there was not a statistically significant difference for scenario one, $F(1,164) = .352, p = .55$; scenario two, $F(1,164) = .344, p = .56$; scenario three, $F(1,164) = 1.121, p = .29$; or scenario four, $F(1,164) = 1.135, p = .29$. An educator's gender did not impact how harshly male and female students were disciplined. Sample sizes, means, and standard deviation are presented in Tables 4.32 through 4.35 for each scenario.

Table 4.32

Sample, Mean and Standard Deviation by Gender of Educator and Student Gender for Scenario 1

	Participant Gender					
	Male			Female		
Student Gender	n	Mean	SD	n	Mean	SD
Male	23	3.09	.793	18	2.94	.802
Female	70	2.87	.850	57	2.89	.646

Table 4.33

Sample, Mean and Standard Deviation by Gender of Educator and Student Gender for Scenario 2

	Participant Gender					
	Male			Female		
Student Gender	n	Mean	SD	n	Mean	SD
Male	23	3.87	.626	18	3.83	.857
Female	70	3.53	.793	57	3.65	.668

Table 4.34

Sample, Mean and Standard Deviation by Gender of Educator and Student Gender for Scenario 3

	<u>Participant Gender</u>					
	Male			Female		
Student Gender	n	Mean	SD	n	Mean	SD
Male	23	4.09	.848	18	3.78	.878
Female	70	3.87	.883	57	3.88	.709

Table 4.35

Sample, Mean and Standard Deviation by Gender of Educator and Student Gender for Scenario 4

	<u>Participant Gender</u>					
	Male			Female		
Student Gender	n	Mean	SD	n	Mean	SD
Male	23	3.61	1.08	18	3.50	.924
Female	70	3.17	.964	57	3.44	1.01

To increase understanding of results and to provide a quick reference for readers, Table 4.36 outlines the statistically significant results for question eleven.

Table 4.36

Summary of Statistically Significant Differences for Question Eleven

Scenario 1: No statistically significant differences
Scenario 2: No statistically significant differences
Scenario 3: No statistically significant differences
Scenario 4: No statistically significant differences

Table 4.37 provides a summary of the findings from the scenarios, statistical test used and effect size. Appendix G contains a figure summarizing variables that were considered and if a statistically significant difference existed or not.

Table 4.37

Summary of Scenario Statistically Significant and Meaningful Results Interpretation Matrix

Scenario	Finding	Statistical Test	Effect Size with Interpretation
2	Male educators disciplined light skin students more "harshly" than did female educators.	Analysis of Variance	Medium $\eta^2 = .07$
3	Light skin males are disciplined more "harshly" than dark skin males.	Analysis of Variance	Medium $\eta^2 = .08$
4	Medium skin female students are disciplined more "harshly" than white female students.	Analysis of Variance	Medium $\eta^2 = .09$

Chapter 5: Discussion, Conclusion, and Recommendations

Discussion of Findings

I began this journey wanting to understand the school discipline gap better, specifically wanting to understand the interracial discipline gap identified in the literature review (Hannon, Defina & Bruch, 2013; Blake et al., 2017). I knew I wanted to create a study that was unique and would provide meaningful discourse about an understudied area. This study was developed based upon what was in the research literature about colorism to explore a topic that is not comprehensively studied or discussed outside of the African American community. Based on what research has been done as well as anecdotal stories in the African American community, my findings puzzled me. At the same time, they provide additional information on colorism and suggest additional research related to how discipline decisions are made.

While the student pictures, scenarios, and the surveys were piloted, in retrospect, I think that changes in the way the scenarios were rated, the scenarios themselves, and the pictures used might need further modifications. I was puzzled by the inconsistencies in responses across the scenarios and skin tone groups. While future studies with the modifications I suggest, might result in the same findings. I would hope such studies might be undertaken for comparison and a clearer understanding of colorism.

In examining differences in discipline by skin tone of the student, as well as gender and race of teacher, I found twelve statistically significant results, three were meaningful, accounting for 7% or more of the variance. No one scenario consistently provided statistically significant or meaningful results. The scenarios themselves produced scattered and unpredictable results. Additionally, my findings did not echo the conclusions presented in the literature review.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Because of the scattered findings, it is not possible to state that a student's skin tone does or does not play a part in how "harshly" educators discipline students. It is possible that the design of this study may account for the lack of consistently meaningful results that align with the literature review. Given the inconsistent findings, I considered four explanations for the discipline responses across scenarios: the content of the scenarios, order of scenarios, and rating scale. The content of the scenarios were created from real classroom discipline infractions and by using the top three infractions identified by the VDOE as guidance. There was one statistically significant and meaningful results for every scenario except for scenario one. However, I am unable to conclude if the scenarios themselves played a role in the findings of this study or some other factor played a role because of a methodological flaw on my part. The order of the scenarios could have inadvertently impacted student discipline ratings. Although the scenarios were piloted and then revised, aside from putting the practice question first, the order of the scenarios was not considered during the piloting phase. Additionally, my study asked participants to rate how likely they were to assign each possible discipline decision (suspend, office referral, verbally reprimand, call parents, ignore) to a given scenario. Instead of presenting five questions about the likelihood of selecting a specific decision, there may have more significant and meaningful results if participants were asked to select one of the four possible decisions to assign to the discipline scenario. Findings might have been more consistent if participants were given fewer options to select from.

In making sense of my findings, I examined the intersection between student gender and skin tone. When looking at how gender and skin tone impacted discipline decisions, I found similar, but not identical, outcomes as the literature. Hannon, DeFina, and Burch (2013) and Blake et al., 2017 found that having a darker skin tone significantly increased the odds of

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

suspension for African American Adolescents (both male and female students). In scenario four, medium skin tone female students were disciplined more “harshly” than white female students. I don't have an explanation for why medium skinned toned female students, which was not the darkest skin tone in the study, was rated more harshly than white female students. More research is needed to better understand the results.

In scenario three, light skin males were disciplined more “harshly” than dark skin males. This finding was not noted in the literature. Based on the literature, I was expecting to find that dark skin males were disciplined more “harshly” than lighter skin males, since this finding was suggested throughout the literature review. Additionally, in scenario two, my study found that male educators disciplined light skin students as a whole, more “harshly” than female educators. While I cannot definitively explain why the findings in scenarios two and three occurred, I explored a possible reason. From experience and by using a commonsense approach, one way to make sense of the results is that light skin students may be expected to behave better than dark skin students. Thus, when a light skin male misbehaves, he is disciplined more “harshly” than the dark skin male because of this expectation.

Although no mention of student skin tone or gender, Marucci (2019) found that teachers were harsher against white students, which contradicts the anti-Black implicit bias that is implied in the literature about school discipline. Results from Marucci's study suggest that social desirability may be at play. Social desirability attempts to override anti-Black bias, which could harmfully lower the expectations teachers have for black students. Social desirability could help explain why the light skin male students were disciplined more “harshly” than the dark-skinned male students. Perhaps social desirability was not a factor in Hannon et al., (2013) and Blake et al., (2017) because these studies used pre-existing, longitudinal data

sets which would allow educators behaviors to be measured without the threat of social desirability or even the Hawthorne effect. However, my study and Marucci (2019) used vignettes to understand discipline decisions which may increase social desirability and the Hawthorne effect.

Racial Mechanics Model: CRT and Implicit Social Cognition Theory

Implicit beliefs about a group of people or people of various hues is formed because of historical stereotypes perpetuated in the media and society as a whole. Critical race theorists argue that racism controls the political, social, and economic aspects of American society, the field of education falls under all three arenas. The Racial Mechanics Model, explains that once a "perceiver" sees the "target," beliefs are formed simultaneously and implicitly, based on racial categories. Those categories then take on racial meanings, which alters how the "perceiver" interacts with the "target." In the case of this study, the perceiver is the educator, and the target is the student.

Although the results of my study are inconclusive and leave many questions unanswered, it is important to acknowledge the findings discussed in the literature review that highlight the intraracial school discipline gap. Therefore, although I cannot conclude from this study that a skin tone bias impacts an educator's discipline decision, I am still left with a body of research that indicates this to be true. It is crucial to make sense of what is currently known from the literature review. The interracial and intraracial discipline gap is complex and multi-dimensional and requires further studies and investigation regardless of my findings. Implicit social cognition theory and CRT are helpful tools for understanding the body of research that does indicate both race and color bias (whether implicit or explicit) are prevalent in student discipline decisions.

CRT provides a framework to analyze racial disparities using five tenants: counter-storytelling; the permanence of racism; whiteness as property; interest conversion; and critique of liberalism. Although all five tenants help explain racial disparities, I will focus on how permanence of racism, whiteness as property, and critique of liberalism can help explain the findings in the literature. I will later discuss the importance of using counter-story telling in giving a voice to students and educators as it relates to race and colorism.

Permanence of racism in CRT suggests that racism is an inherent part of American society, which privileges white individuals over people of color. Policies put in place at the school level have the ability to disproportionately impact one group of students over another, negatively or positively. The discipline gap is the result of school discipline policies that place students of color at higher risk for being over punished through suspension and expulsion when compared to their white counterparts. The scenarios used in my study were based on the most cited infractions, which were mostly subjective. A system that allows the practice of suspending students for subjective infractions allows the discipline gap to persist. Educators must acknowledge that racism and skin tone bias exists due to the historical foundation in the United States and that such bias (implicitly or explicitly) may be at play when an educator decides how to handle/manage a student's behavior.

CRT describes whiteness as property as the notion that white people have a right of possession, the right to employment, the right of exclusion, and the right to use and enjoyment. Whiteness serves not only as identity and status but also as a property right (Bell, 1995; Harris, 1993) that extends beyond an individual or group of White people to include institutional structures such as education, politics, religion, the presidency of the United States which have been historically occupied by Whites. Although decades since *Brown vs. Board of Education*,

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

education is still a notion that is reserved for white people. Using this theory to gain understanding, it is possible that educators of all races unknowingly or knowingly believe that white students have a right to a certain standard of education that other students may not and therefore cannot afford to miss instructional time due to poor behavior. None of the findings in my study suggests that white students or the control group were disciplined more "harshly" than students of color.

In line with whiteness as property, white innocence is historically predicated on the criminalization and violation of (primarily) black bodies, the framing of whites throughout the history of the United States as inherently innocent and blacks as guilty (Fylkesnes, 2018). White innocence is a theory that can explain how educators of all races view white students. Educators may be willing to give white students additional opportunities to correct or modify their behaviors. They may view poor behaviors from white students as more innocent or "child-like" in nature (Sullivan, 2014). Additionally, white innocence can be used to explain how educators may view their role in contributing to the discipline gap. Foster (2019) discusses an "enlightenment narrative" and uses this term to identify white student leaders, although it could apply to all educators who are trained in Eurocentric teaching styles. White leaders use the strategy of presenting oneself as racially, conscious and progressive. By appearing "enlightened," these leaders situated themselves as racially enlightened and aware, placing themselves outside of the problem of white supremacy. This same principle may be applied to all educators. It is possible that educators are held to a high standard by society and are viewed as being educated and too "enlightened" to engage in any form of racial bias against students. If educators are viewed as believers of equality, they may knowingly or unknowingly perpetuate and reinforce systemic racism via colorism. While I use this theory of enlightenment to discuss CRT

tenant, whiteness as property, one could argue that it could apply to critique of liberalism as well.

Critique of liberalism is a tenant of CRT that discusses the notion of colorblindness, neutrality of law and equal opportunity for all (Decuir & Dixson, 2004; Hiraldo, 2010). Results from the literature show that "colorblindness" may not be a real concept that is at play during the discipline process. If "colorblindness" existed, I would expect to see all students, regardless of skin tone disciplined the same or at least similarly. "Colorblindness" is a mechanism that allows people to ignore racist policies that perpetuate social inequality (Hiraldo, 2010; Decuir & Dixson, 2004). The notion of colorblindness and the belief that "colorblindness" is an effective tool to deal with racial issues in the United States keeps critical conversations from occurring. Stakeholders must begin to first talk about colorism and understand that racial disparities are not binary (black and white) but are instead multidimensional and layered. Stakeholders must understand continuing the idea of "colorblindness" may influence educator's disciplinary decisions. By seeing all students, but specifically black students as "the same," educators may unknowingly under or over discipline, based on race and also based on hue. Educators cannot be color blind to a student's race or skin tone. Educators cannot discipline or even educate students without using a historical lens to understand the implication of colorism in the United States. Research need to continue to understand how a within- race stratification can impact students of color. It will take work on the part of all educational stakeholders to shift from the notion of "colorblindness" to seeing race not only through a black and white lens but through a lens of acknowledging how students' race and skin tone can impact educational and discipline outcomes.

Current State of Affairs- Policy and Political Climate

It is essential to discuss the interracial and intraracial discipline gap within the context of the current political climate. The school discipline gap is gradually gaining attention across the United States. The need to bring awareness to this issue was recognized with the creation of 2014 Dear Colleague Letter on the Nondiscriminatory Administration of School Discipline by the U.S. Department of Education and the U.S. Department of Justice, under the guidance of the Office of Civil Rights during the presidency of Barack Obama. This document offered guidance to parents, advocates, educators, and school divisions about approaches to administering student discipline without discriminating by race, ethnicity, or gender. The document detailed how discrimination against students is prohibited in the United States due to Titles IV and VI and discussed how stakeholders could follow a "Different Treatment" flowchart to determine if a school district has intentionally or unintentionally discriminated against a student. If a school was found to have discriminated against a student, stakeholders were able to follow procedures to begin a federal investigation process.

However, since the creation of this document to move educators toward a more complete understanding of implicit bias as well as to provide fairer discipline practices, presidential leadership in the United States has changed, and with the changes, the political climate has taken on a new set of values and priorities. On December 21, 2018, the Commission on School Safety, under the guidance of President Donald Trump, decided to withdraw statements of federal policy and guidance so that individual states and local school divisions could play a primary role in establishing appropriate discipline policies and practices. If discriminatory practices are found to have occurred, school divisions will not lose federal funding. Withdrawing these statements places the most vulnerable population (children, families living in poverty) at risk for receiving

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

an inequitable education, providing no guidance on issuing complaints against school divisions when they believe a child is being unfairly disciplined. Now states and localities are in control of governing disciplinary practices.

It is not certain if racist behaviors have increased or if technological advances (camera phones) make documenting these behaviors easier. Edwards and Rushin (2018) reported a "Trump Effect" to explain the rise in reported hate crimes documented across the United States. The "Trump Effect" states that it, not just the remarks made by Donald Trump throughout his political campaign, but that his election validated racist notions that caused a surge in hate crimes. Although this study did not specifically analyze educator hate crimes or racist incidents, it is possible that the "Trump Effect has also impacted educators." It is possible that educators who already held racist ideals are more comfortable expressing their ideas in and outside of the classroom now that the leader of the country has expressed similar opinions. Educators may feel supported by the current leader and therefore feel as though their beliefs can be freely expressed in the classroom. There has been an increase in publicized situations surrounding educator's use of racist hate speech and affiliation with hate groups (Edwards and Rushin, 2018). For example, in March 2018, Dayanna Volitich, a middle school teacher in Florida was accused of hosting a white nationalist podcast (Cable News Network, 2018). On one of many podcasts, Volitich argued that science proves that certain races are smarter than others and bragged about teaching white nationalism in her classroom.

In March 2019, in Houston, Texas, a school teacher was videographed verbally harassing a Muslim woman outside of a supermarket (*Houston Chronicle*, 2019). That same month, in Richmond, Virginia, a school resource officer was fired for being a part of a racist organization (*Richmond Times-Dispatch*, 2019). These examples remind us that explicit racial

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

bias is present in schools and society, and possibly that implicit racial bias is at play when disciplining students of color. Teachers may carry their beliefs with them, and either consciously or unconsciously weave them into instruction and daily interactions with students and community members. Educators may unintentionally demonstrate bias or racist ideologies as with the many cases of educators teaching lessons that are not culturally sensitive and perpetuates stereotypes and racist ideals.

Conclusion

I began this study with the assumption that most teachers do not intentionally hold an explicit bias against students but do so implicitly out of ignorance. Based on findings from my study, I cannot conclude that educators intentionally or unintentionally hold a bias, a survey – even one with a simulated case -- can never capture all the complexities that are at play in the classroom. I relied on findings presented in the literature review to determine that colorism in education, as it related to student discipline outcomes, is an area worth continued exploration. Using implicit bias as the theoretical framework, I entered into this study accepting that educators may unknowingly engage in discipline practices that disproportionately impact students of color. CRT proposed that we are not color blind or race-neutral nor objective. Discussing colorism in education parallels the discussion of racism in education. Regardless of the results from this study, research on colorism provides strong evidence colorism still exists and that stakeholders have to keep in mind the historical nature of colorism within the United States, when evaluating school practices.

Race is a nuanced social construct that goes beyond black and white. Educators need to expand the dialogue on race and power to include an understanding of colorism, which the literature shows that it is embedded in American culture because of the history of this country.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

The intraracial and interracial discipline gap described in the literature continues white supremacy by ignoring that educators are a part of a larger society who hold biases and stereotypes. Additional research is needed to understand if teachers of color hold white supremacist view when interacting with students of color. Because of the possibility that teachers regardless of race engage in implicit bias, all teachers must develop an awareness of the interracial and intraracial student discipline gap by seeing racism as a continuum and not a defined black and white issue. Additionally, stakeholders need to understand the legal ramifications of implicitly or explicitly disciplining students based on skin tone is a civil rights issue.

Recommendations

This study is a small contribution to the field of education. At the very least, I have examined colorism and the possibility of skin tone bias in discipline. Based on the literature, colorism is understudied and yet maybe so crucial in understanding and combating the discipline gap. Research needs to continue to clarify how to prepare educators to work with an increasingly diverse student body. Future research should study discipline trends to include skin tone and consider adding a skin tone score when infractions are reported to the state level Departments of Education.

My study did not find consistent patterns, and I cannot conclude that a student's skin tone does or does not influence discipline decisions. Given the impact skin tone has been shown to have on other areas of life, it is a topic that needs further exploration. More research is needed to understand how the intersection of educator race and gender impact discipline decisions. Also missing from the literature are studies of the intersection of race and gender, both for educators who are making discipline decisions and students who are the focus of the

decisions. Based on the findings in the literature review and my study, I make the following recommendations for research, practice, and policy.

1) Professional Development

School divisions need to provide ongoing professional development opportunities pertaining to diversity and racial and skin tone bias. Without proper training, educators may spend their careers engaging in implicit bias against specific types of students. It is assumed that educators enter the field of public k-12 education because they enjoy working with children, enjoying teaching their specific content area and want students to obtain new knowledge. Additionally, educators need training on how to respond to an increasingly diverse student body. Educators may react differently to students during the discipline process, depending on how much experience they have interacting with students that are not of their race. Educators should be supported in pursuing their passion while also understanding social justice and equity issues at play in the educational setting.

2) Office of Equity: Compliance, Data Collection, and Monitoring

School divisions have to create an office strictly dedicated to ensuring equitable treatment of students. Having a dedicated office space and personnel to maintain the office will ensure that teachers, students, and parents have a direct point of contact when they suspect a student is being discriminated against due to race, skin tone, sexual orientation and even religion. In addition, school divisions need to conduct equity audits that not only address the school discipline gap but include student skin tone in suspension data. Instead of only recording racial discipline data, skin tone data should also be collected. An equity audit is a tool that can be used to gain insight on an identified discrepancy, like school discipline in this case. This office of equity would be responsible for completing equity audits and frequent and ongoing data

monitoring.

3) Qualitative Research

More research is needed to understand the concept of implicit bias in the field of education as it pertains to colorism. Future research should utilize qualitative data to better understand skin tone bias in the actual words of educators, students, and parents. For example, I would be interested in understanding the educational experiences of black educators as it relates to colorism and understand if the experiences (negative or positive) of black high school students are attributed to skin tone.

A qualitative study about this topic would allow an in-depth analysis of the actual experiences of educators and students as it relates to their experiences with colorism. More research is needed to understand colorism as it relates to the educational experiences of students who are racially ambiguous or bi-racial. Keeping with the CRT tenant of storytelling, the experiences of students and educators are mostly narrative. Georgetown Law Center demonstrated how compelling the lived experiences of people of color could be with its development of *Girlhood Interrupted*, which enables black women to share their experiences of being a black girl in the school setting. I believe quantitative data cannot honestly answer all questions related to this topic or explain the lived experiences of teachers and students.

References

- Action partnership on interventions for black children exposed to violence and victimization. (2014). Retrieved February 02, 2016, from <https://victimsofcrime.org/our-programs/other-projects/youth-initiative/interventions-for-black-children's-exposure-to-violence/black-children-exposed-to-violence>
- Annamma, S., Anyon, Y., Joseph, N., Farrar, J., Greer, E., Downing, B., & Simmons, J. (2019). Black girls and school Discipline: The complexities of being overrepresented and understudied. *Urban Education, 54*(2), 211-242.
- Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38*(2), 113-125.
- Balfanz, R., Byrnes, V., & Fox, J. Sent Home and Put Off-Track: The Antecedents, Disproportionalities, and Consequences of Being Suspended in Ninth Grade.[Internet]. Everyone Graduates Center, School of Education, Johns Hopkins.
- Bates, L. A., & Glick, J. E. (2013). Does it matter if teachers and schools match the student? Racial and ethnic disparities in problem behaviors. *Social Science Research, 42*(5), 1180-1190.
- Beattie, G., Cohen, D., & McGuire, L. (2013). An exploration of possible unconscious ethnic biases in higher education: the role of implicit attitudes on selection for university posts. *Semiotica (197)*, 171-201.
- Ben-Zeev, A., Dennehy, T. C., Goodrich, R. I., Kolarik, B. S., & Geisler, M. W. (2014). When an “educated” Black man becomes lighter in the mind’s eye: Evidence for a skin tone memory bias. *Sage Open, 4*(1), 2158244013516770.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

- Berger, D. G. (1969). They all look alike. unpublished doctoral dissertation, Vanderbilt University.
- Bertrand, M., & Mullainathan, S. (2004). Are emily and greg more employable than lakisha and jamal? A field experiment on labor market discrimination. *The American Economic Review*, 94(4), 991-1013.
- Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological science*, 15(10), 674-679.
- Blair, I. V., Chapleau, K. M., & Judd, C. M. (2005). The use of Afrocentric features as cues for judgment in the presence of diagnostic information. *European Journal of Social Psychology*, 35(1), 59-68.
- Blake, J. J., Butler, B. R., Lewis, C. W., & Darenbourg, A. (2011). Unmasking the inequitable discipline experiences of urban black girls: Implications for urban educational stakeholders. *The Urban Review*, 43(1), 90-106.
- Blake, J. J., Keith, V. M., Luo, W., Le, H., & Salter, P. (2017). The role of colorism in explaining African American females' suspension risk. *School Psychology Quarterly*, 32(1), 118.
- Branigan, A. R., Wildeman, C., Freese, J., & Kiefe, C. I. (2017). Complicating colorism: Race, skin color, and the likelihood of arrest. *Socius*, 3, 2378023117725611.
- Bristol, T., & Riehl, Carolyn. (2014). Black men of the classroom: An exploration of how the organizational conditions, characteristics, and dynamics in schools affect black male teachers' pathways into the profession, experiences, and retention, ProQuest Dissertations and Theses.
- Brown, T.D., Dane F.C., & Durham MD. (1998) Perception of race and ethnicity. *Journal of*

Social Behavior & Personality, 13(2), 295-306.

Burke, M. (2008). Colorism. *International encyclopedia of the social sciences* (Vol. 2., pp. 17 – 18). Detroit, MI: Thomson Gale.

Bukach, C. M., Cottle, J., Ubiwa, J., & Miller, J. (2012). Individuation experience predicts other-race effects in holistic processing for both caucasian and black participants. *Cognition*, 123(2), 319-324.

Carr, E., Dunlap, G., Horner, R., Koegel, R., Turnbull, A., Sailor, W., Anderson, J., Albin, R., Koegel, L., & Fox, L. (2002). Positive behavior support: Evolution of an applied science. *Journal of Positive Behavior Interventions*, 24(4), 4-16.

Carpusor, A. G., & Loges, W. E. (2006). Rental discrimination and ethnicity in names. *Journal of Applied Social Psychology*, 36(4), 934-952.

Chao, G. T., & Willaby, H. W. (2007). International employment discrimination and implicit social cognition: New directions for theory and research. *Applied Psychology*, 56(4), 678-688.

Cleeremans, A. (2006). Implicit learning models. *Encyclopedia of cognitive science*. John Wiley & Sons, Ltd. doi:10.1002/0470018860.s00097

Coard, S. I., Breland, A. M., & Raskin, P. (2001). Perceptions of and preferences for skin color, black racial identity, and self-esteem among african americans. *Journal of Applied Social Psychology*, 31(11), 2256-2274.

Clair, M., & Denis, T. (2015). "Sociology of Racism." In *International Encyclopedia of Social and Behavioral Sciences*. Vol. 19, 2nd ed., edited by James D. Wright, 857–863. London: Elsevier.

Clark, K. B., & Clark, M. P. (1950). Emotional factors in racial identification and preference in

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

- Negro children. *The Journal of Negro Education*, 19(3), 341-350.
- Cooper, A. (2010). CNN pilot demonstration.
- Criminal Justice Fact Sheet. (2014). Retrieved December 20, 2015, from <http://www.naacp.org/pages/criminal-justice-fact-sheet>
- Darensbourg, A., Perez, E., and Blake, J. (2010). Over representation of african american males in exclusionary. *Journal of African American Males in Education*, 1(3).
- DeCuir, J. T., & Dixson, A. D. (2004). "So when it comes out, they aren't that surprised that it is there": Using critical race theory as a tool of analysis of race and racism in education. *Educational researcher*, 33(5), 26-31.
- Delgado, R., & Stefancic, J. (2017). *Critical race theory: An introduction* (Third ed., Critical America)
- DeMatthews, D. E., Carey, R. L., 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.
- Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V., & Johnson, S. L. (2006). Looking deathworthy: Perceived stereotypicality of black defendants predicts capital- sentencing outcomes. *Psychological Science*, 17(5), 383-386.
- Ellis, H. D., Deregowski, J. B., & Shepherd, J. W. (1975). Description of white and black faces by white and black subjects. *International Journal of Psychology*, 10(2), 119-123.
- Farrell, W. C., & Olson, J. L. (1983). Kenneth and mamie clark revisited: Racial identification and racial preference in dark-skinned and light-skinned black children. *Urban Education*, 18(3), 284-297.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

- Fedders, B. (2009). Race and market values in domestic infant adoption. *NCL Rev.*, 88, 1687.
- Ferguson, A. A. (2001). *Bad boys: Public schools in the making of black masculinity*. University of Michigan Press.
- Foster, P. (2019). *Policy and practice in multicultural and anti-racist education: A case study of a multi-ethnic comprehensive school*. Routledge.
- Foucault, M. (2012). *Discipline and punish: The birth of the prison*. Vintage
- Francis, D. V. (2012). Sugar and spice and everything nice? Teacher perceptions of black girls in the classroom. *The Review of Black Political Economy*, 39(3), 311-320.
- Fylkesnes, S. (2018). Whiteness in teacher education research discourses: A review of the use and meaning making of the term cultural diversity. *Teaching and Teacher Education*, 71, 24–33.
- Gibbons, A. (2006). *The first human: The race to discover our earliest ancestors* (1st ed.). New York: New York: Doubleday.
- Greenwald, A. G., & Krieger, L. H. (2006). Implicit bias: Scientific foundations. *California Law Review*, 94(4), 945-967.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological review*, 102(1), 4.
- Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher*, 39(1), 59-68.
- Gyimah-Brempong, K., & Price, G. N. (2006). Crime and punishment: And skin hue too? *American Economic Review*, 96(2), 246-250.
- Hagiwara, N., Kashy, D. A., & Cesario, J. (2012). The independent effects of skin tone and facial features on Whites' affective reactions to Blacks. *Journal of Experimental Social Psychology*, 48(4), 892-898.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Hall, R. E. (1993). Clowns, buffoons, and gladiators: Media portrayals of African- American men. *The Journal of Men's Studies, 1*, 239-251.

Hall, R. E. (2005). From the psychology of race to the issue of skin color for people of African descent. *Journal of Applied Social Psychology, 35*, 1958—1967.

Hall, R. E. (2016). The Bleaching Syndrome: The Role of Educational Intervention. *Theory Into Practice, 55*(1), 62-68.

Hamilton, D., Goldsmith, A. H., & Darity, W. (2009). Shedding “light” on marriage: The influence of skin shade on marriage for black females. *Journal of Economic Behavior and Organization, 72*(1), 30-50.

Hannon, L. (2015). White colorism. *Social Currents, 2*(1), 13-21.

Hannon, L., & DeFina, R. (2016). Reliability Concerns in Measuring Respondent Skin Tone by Interviewer Observation. *Public Opinion Quarterly, 80*(2), 534-541.

Hannon, L., DeFina, R., & Bruch, S. (2013). The Relationship Between Skin Tone and School Suspension for African Americans. *Race and Social Problems, 5*(4), 281-295.

Harrison, M. S., & Thomas, K. M. (2009). The hidden prejudice in selection: A research investigation on skin color bias. *Journal of Applied Social Psychology, 39*(1), 134-168.

Heilbrun, A., Cornell, D., & Lovegrove, P. (2015). Principal attitudes regarding zero tolerance and racial disparities in school suspensions. *Psychology in the Schools, 52*(5), 489-499.

Heitzeg, Nancy A.(2009). *Education or incarceration: zero tolerance policies and the school to prison pipeline*. Forum on Public Policy. Winter.

Hills, P. J., & Lewis, M. B. (2006). Reducing the own-race bias in face recognition by shifting

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

- attention. *The Quarterly Journal of Experimental Psychology*, 59(6), 996-1002.
- Hines-Datiri, D., & Carter Andrews, D. (2017). The effects of zero tolerance policies on black girls: Using critical race feminism and figured worlds to examine school Discipline. *Urban Education*, 004208591769020.
- Hiraldo, P. (2010). The role of critical race theory in higher education. *The Vermont Connection*, 31(1), 7.
- Hunter, M. (2016). Colorism in the Classroom: How skin tone stratifies African American and Latina/o students. *Theory Into Practice*, 55(1), 54-61.
- Hunter, M. L. (2005). *Race, gender, and politics of skin tone*. New York: Routledge.
- Irby, D. J. (2014). Revealing racial purity ideology fear of Black–White intimacy as a framework for understanding school discipline in post-brown schools. *Educational Administration Quarterly*, 50(5), 783-795.
- Kang, J. (2005). Trojan horses of races. *Harvard Law Review*, 118(5), 1489-1593.
- Kennedy, J. F. (. F., 1917-1963. (1964). *A nation of immigrants* (Rev. and enl. ed. Introd. by Robert F. Kennedy. ed.). New York: New York : Harper and Row.
- Kinsler, J. (2011). Understanding the black–white school discipline gap. *Economics of Education Review*, 30(6), 1370-1383.
- Klonoff, E., & Landrie, H. (2000). Is skin color a marker for racial discrimination? Explaining the skin color–hypertension relationship. *Journal of Behavioral Medicine*, 23(4), 329-338.
- Kunesh, C. E., & Noltemeyer, A. (2019). Understanding disciplinary disproportionality: Stereotypes shape pre-service teachers' beliefs about black boys' behavior. *Urban Education*, 54(4), 471-498.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Ladson-Billing, G. Tate iv, WF 1995. *Towards a Critical Race Theory of Education*. Columbia University, *Teacher college*, 97(1), 49-68.

Lindsay, C. A., & Hart, C. (2017). Teacher Race and School Discipline. *Education Next*, 17(1), 1-9.

Maddox, K. B.(2004). Perspectives on racial phenotypicality bias. *Personality and Social Psychology Review*, 8, 383-401.

Marira, T. D., & Mitra, P. (2013). Colorism: Ubiquitous yet understudied. *Industrial and Organizational Psychology*, 6(1), 103-107.

Marcucci, O. (2019). Implicit bias in the era of social desirability: Understanding antiblackness in rehabilitative and punitive school discipline. *The Urban Review*, 52(1), 47-74.

McGee, E. O., Alvarez, A., & Milner IV, H. R. (2016). Colorism as a salient space for understanding in teacher preparation. *Theory into Practice*, 55(1), 69-79.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, 7(1), 3.

Milner, H. R. (2013). Why are students of color (still) punished more severely and frequently than white students? *Urban Education*, 48(4), 483-489.

Monroe, C. R. (2013). Colorizing educational research african american life and schooling as an exemplar. *Educational Researcher*, 42(1), 9-19.

Morris, E. W. (2007). "Ladies" or "loudies"? Perceptions and experiences of black girls in classrooms. *Youth & Society*, 38(4), 490-515.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

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

Norwood, K. J. (Ed.). (2014). *Color matters: Skin tone bias and the myth of a postracial America*. Routledge.

Office of Civil Rights Data Collection. *Data snapshot: School discipline*. (2014).

Retrieved January 5, 2015, from <http://blogs.edweek.org/edweek/rulesforengagement/CRDCSchoolDisciplineSnapshot.pdf>

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

Pigott, R. L., & Cowen, E. L. (2000). Teacher race, child race, racial congruence, and teacher ratings of children's school adjustment. *Journal of School Psychology, 38*(2), 177-195.

Rachlinski, J. J., Johnson, S. L., Wistrich, A. J., & Guthrie, C. (2009). Does unconscious racial bias affect trial judges? *Notre Dame Law Review, 84*(3), 1195-1246.

Ragsdale, L. A. (2013). *An Analysis of Three White Male High School Assistant Principal Perceptions of Black Males Students in a Suburban District* (Doctoral dissertation).

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

Rich, M. (2015). *Where are the teachers of color?* New York Times. Retrieved December 12, 2015, from <http://www.nytimes.com/2015/04/12/sunday-review/where-are-the-teachers-of-color.html>

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

- Ryabov, I. (2013). Colorism and school-to-work and school-to-college transitions of African American adolescents. *Race and Social Problems*, 5(1), 15-27.
- Secord, P. F., Bevan, W., & Katz, B. (1956). The Negro stereotype and perceptual accentuation. *The Journal of Abnormal and Social Psychology*, 53(1), 78.
- Sellers, R., Smith, M., Shelton, N., Rowley, S., & Chavous, T. (1998). Multidimensional model of racial identity: A reconceptualization of african american racial identity. *Personality and Social Psychology Review*, 2(1), 18-39.
- Shepherd, J. W., Deregowski, J. B., & Ellis, H. D. (1974). A cross-cultural study of recognition memory for faces. *International Journal of Psychology*, 9(3), 205-212.
- Siegel, E. F., Dougherty, M. R., & Huber, D. E. (2012). Manipulating the role of cognitive control while taking the implicit association test. *Journal of Experimental Social Psychology*, 48(5), 1057-1068.
- Skiba, R. J., & Knesting, K. (2002). Zero tolerance, zero evidence: An analysis of school disciplinary practice. In R. J. Skiba & G. G. Noam (Eds.), *New directions for youth development (no. 92: Zero tolerance: Can suspension and expulsion keep schools safe?)*(pp. 17–43). San Francisco: Jossey- Bass.
- Skiba, R., & Edl, H. (2004). The Disciplinary Practices Survey: How Do Indiana's Principals Feel about Discipline. Children Left Behind Policy Briefs. Supplementary Analysis 2-C. *Center for Evaluation and Education Policy, Indiana University*.
- Skiba, R. J., Horner, R. H., Chung, C., Rausch, M. K., May, S. L., & Tobin, T (2011). Race is not neutral: A national investigation of african american and latino disproportionality in school discipline. *School Psychology Review*, 40(1), 85-107.

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Smith-McLallen, A., Johnson, B. T., Dovidio, J. F., & Pearson, A. R. (2006). Black and white: The role of color bias in implicit race bias. *Social Cognition, 24*(1), 46-73.

Staats, C.(2014). *Implicit racial bias and school discipline disparities*. Kirwan Institute, 1-20. Retrieved July 17, 2014, from <http://kirwaninstitute.osu.edu/wp-content/uploads/2014/05/ki-ib-argument-piece03.pdf>

Steele, C.M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of african-Americans. *Journal of Personality and Social Psychology, 69*, 797-811.

Stepanova, E. V., & Strube, M. J. (2009). Making of a face: Role of facial physiognomy, skin tone, and color presentation mode in evaluations of racial typicality. *The Journal of Social Psychology, 149*, 66–81.

Stepanova, E. V., & Strube, M. J. (2012). The role of skin color and facial physiognomy in racial categorization: Moderation by implicit racial attitudes. *Journal of Experimental Social Psychology, 48*(4), 867-878.

Sullivan, S. (2014). *Good White people: The problem with middle-class White anti-racism*. Albany, NY: SUNY Press.

The 2017 brown center report on american education: How well are american students learning?(pp. 1-33). (2017). Washington, D.C: The Brookings Institution.
doi:<https://www.brookings.edu/wp-content/uploads/2017/03/2017-brown-center-report-on-american-education.pdf>

The schott 50 state report on public education and black males. (2015, February 11). Retrieved February 2, 2016, from <http://www.blackboysreport.org/2015-black-boys-report.pdf>

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Thompson, M. S., & McDonald, S. (2016). Race, skin tone, and educational achievement. *Sociological Perspectives*, 59(1), 91-111.

U.S. Department of Education and U.S Department of Justice. Joint - Dear Colleague Letter. (2014, January 8). Retrieved October 1, 2014, from <http://www2.ed.gov/about/offices/list/ocr/letters/colleague-201401-title-vi.html>

Viglione J., Hannon L., DeFina R.(2011). The impact of light-skin on prison time for Black female offenders. *Social Science Journal*, 48, 250-258.

Villegas, A. M., Strom, K., & Lucas, T. (2012). Closing the racial/ethnic gap between students of color and their teachers: An elusive goal. *Equity & Excellence in Education*, 45(2), 283-301.

Walker, P. M., & Hewstone, M. (2006). A developmental investigation of other-race contact and the own-race face effect. *British Journal of Developmental Psychology*, 24(3), 451-463.

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

Washington Research Project Defense Fund. (1975). *School suspensions: Are they helping children?:A report*. Washington, D.C. Washington Research Project.

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

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

Zero tolerance and alternative strategies: A fact sheet for educators and policymakers. (2014, January 1). NASP Fact Sheet. Retrieved July 23, 2014, from http://www.nasponline.org/resources/factsheets/zt_fs.aspx

APPENDIX A

Survey Questions

Hello:

You are invited to participate in a research survey about classroom behaviors.

1) You will have fifteen minutes to complete this survey. Your participation in this study is completely voluntary. Your survey responses are confidential and data from this research will be reported only in the aggregate. There are no foreseeable risks associated with this project.

2) If you feel uncomfortable answering any questions, you can withdraw from the survey at any point. This survey is for educational purposes only. Reproduction or dissemination of this survey for any purpose is prohibited without permission.

3) Thank you for your time and support. Please start the survey now by clicking on the continue button below. If you have questions at any time about the survey or the procedures, you may contact Kierstyn Johnson-Wigfall by email at johnsonkk2@mymail.vcu.edu.

- I agree
- I disagree

1. Are you an educator who works or has worked in the K-12 setting in the United States?

- Yes
- No

2. Which of the following most accurately describes your current position:

- A. Teacher
- B. Principal
- C. Assistant Principal
- D. Support Staff
- E. Other Administrator
- F. Other _____

4. You are now going to read brief scenarios about classroom behaviors exhibited by students.

- 1) After each scenario you will be asked to respond to questions.
- 2) The scenarios take place at United States High School.
- 3) You are teaching a core subject area and you have disciplined this student before for the same behaviors.

5. You are preparing your class to take a major state-wide test. Over the last few weeks, you have constantly reminded Student A not to interrupt others while they are taking the test. Student A is often impatient and constantly calls out in class. This behavior is annoying to you and to many of the other students in your class. You frequently remind Student A that these behaviors

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

are frustrating, but the calling out and interrupting persists.

1. How likely are you to suspend this student?

- Very Unlikely
- Unlikely
- Undecided
- Likely
- Very Likely

2. How likely are you to issue this student an office referral?

- Very Unlikely
- Unlikely
- Undecided
- Likely
- Very Likely

3. How likely are you to verbally reprimand this student?

1. Very Unlikely
2. Unlikely
3. Undecided
4. Likely
5. Very Likely

4. How likely are you to call the student's parents?

- Very Unlikely
- Unlikely
- Undecided
- Likely
- Very Likely

5. How likely are you to ignore the student's behavior?

- Very Unlikely
- Unlikely
- Undecided
- Likely
- Very Likely

6. Over the last few weeks, you have noticed a change in Student A behaviors including constantly lying to you and not being able to complete an entire assignment during the class period. You have sent multiple letters home to Student A's parents about incomplete homework assignments and coming to class without the textbook or notebook. You have also directly addressed excessive lying with Student A on multiple occasions, in which student A responds "It won't happen again".

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Today during class, Student A asked to go to the clinic because of suspicion of a concussion in gym class that was not reported to the gym teacher. You ask Student B to follow Student A to the clinic to ensure safe arrival. Student B reported that Student A went into the restroom, then teacher's workroom and then ran off and could not be found. When Student A showed back up. Student A reported going to the clinic and when questioned Student A admitted to lying to get out of class.

1. How likely are you to suspend this student?
 - Very Unlikely
 - Unlikely
 - Undecided
 - Likely
 - Very Likely

2. How likely are you to issue this student an office referral?
 - Very Unlikely
 - Unlikely
 - Undecided
 - Likely
 - Very Likely

3. How likely are you to verbally reprimand this student?
 - Very Unlikely
 - Unlikely
 - Undecided
 - Likely
 - Very Likely

4. How likely are you to call the student's parents?
 - Very Unlikely
 - Unlikely
 - Undecided
 - Likely
 - Very Likely

5. How likely are you to ignore the student's behavior?
 - Very Unlikely
 - Unlikely

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

- Undecided
- Likely
- Very Likely

7. You are teaching a lesson that students will later be tested on at the end of the quarter. Student A is making paper airplanes and talking non-stop during your lecture. This behavior is common for Student A, and you have spoken to Student A several times, trying to get the student to pay attention in class or at least not disrupt other students.

You tell student A to put the paper airplanes in the trash and then pay attention to your presentation. Student A walks to the trash and throws the first plane away, then returns and gets the second airplane from the desk. Student A then walks that airplane to the trash and throws it away. This continues for two more airplanes. Once seated, student A continues to talk to other students, asking questions and making comments. Classmates appear to be trying to ignore Student A.

Student A continues, using suggestive language and calls one student a “mother-fucker” and responds to a student who asks for the behavior to stop with “Fuck you.”. Throughout this behavior, you have told Student A to sit down and be quiet three times. Finally, Student A sits down but then begins throwing a pencil around the room. Each time, Student A gets up, picks up the pencil, sits down, and then throws it again.

1. How likely are you to suspend this student?
 - Very Unlikely
 - Unlikely
 - Undecided
 - Likely
 - Very Likely

2. How likely are you to issue this student an office referral?
 - Very Unlikely
 - Unlikely
 - Undecided
 - Likely
 - Very Likely

3. How likely are you to verbally reprimand this student?
 - Very Unlikely
 - Unlikely
 - Undecided
 - Likely
 - Very Likely

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

4. How likely are you to call the student's parents?

- Very Unlikely
- Unlikely
- Undecided
- Likely
- Very Likely

5. How likely are you to ignore the student's behavior?

- Very Unlikely
- Unlikely
- Undecided
- Likely
- Very Likely

8. It is Monday morning and you are teaching a new chapter in the class textbook. Student A's head is on the desk and Student A is sleeping. This is not unusual for Student A. Neither is interrupting class. Both things have happened regularly in class and you have talked with Student A several times about the behavior.

You tell student A to "sit up and pay attention". Student A responds to your request by sitting up, but begins to make loud, random animal sounds, such as "quack-quack" and "moo". Some classmates tell Student A to "cut it out". Student A immediately discontinues the noise. However, within minutes, Student A begins to make inappropriate jokes while you are giving directions for homework. Student A then begins making loud farting noises. You turn to Student A and say, "Stop now or you will need to go to the principal's office". Student A replies by mocking you, saying "Stop now or you will have to go to the principal's office."

1. How likely are you to suspend this student?

- Very Unlikely
- Unlikely
- Undecided
- Likely
- Very Likely

2. How likely are you to issue this student an office referral?

- Very Unlikely
- Unlikely
- Undecided
- Likely
- Very Likely

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

3. How likely are you to verbally reprimand this student?
 - Very Unlikely
 - Unlikely
 - Undecided
 - Likely
 - Very Likely

4. How likely are you to call the student's parents?
 - Very Unlikely
 - Unlikely
 - Undecided
 - Likely
 - Very Likely

5. How likely are you to ignore the student's behavior?
 - Very Unlikely
 - Unlikely
 - Undecided
 - Likely
 - Very Likely

9. What is your gender?
 - Male
 - Female

10. Are you Hispanic or Latina/o (A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race.)
 - Yes, Hispanic or Latina/o
 - No, Not Hispanic or Latina/o

11. How would you describe yourself? Select one or more from the following racial groups
 - American Indian or Alaska Native
 - Asian
 - Black or African American
 - Native Hawaii or Other Pacific Islander
 - White
 - Other

12. Teaching Experience: How many years have you taught in the public K-12 setting.
 - Less than 1

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

- 1-3
- 4-6
- 7-9
- 10-12
- 13-15
- More than 15 years

13. Which word below best describes the school where you work?

- Majority low socioeconomic
- Majority middle socioeconomic
- Majority high socioeconomic
- A mixture of socioeconomic
- Other

14. Thank you for taking this survey. If you have questions at any time about the survey or the procedures, you may contact Kierstyn Johnson-Wigfall by email at johnsonkk2@mymail.vcu.edu.

You will now have the opportunity in the space below to give feedback about this survey.

To be entered in the fifty dollar Starbucks gift card drawing:

Email johnsonkk2@mymail.vcu.edu and type "Completed the survey" in the subject line. You will be contacted by email if you are the winner of the drawing.

APPENDIX B

Email Notification

Dear past or current student in the educational leadership program at VCU,

I am a doctoral student at VCU. I am writing to you today seeking your participation in my dissertation research survey. The purpose of my study is to understand classroom behaviors. Data will be analyzed from the perspectives of educators.

Your participation is voluntary. Your responses will remain completely confidential. Participants who complete this survey will have a chance to be entered into a raffle to receive a \$50.00 Starbucks gift card.

Informed consent is presented on the first page of the survey. To participate, please click on the following link: (randomly assigned link inserted here). You will have fifteen minutes to complete this survey once clicking the link.

If you have questions at any time about the survey or the procedures, you may contact Kierstyn Johnson-Wigfall by email at johnsonkk2@mymail.vcu.edu. Your participation in this study is greatly appreciated.

Thanks,

Kierstyn Johnson-Wigfall
Virginia Commonwealth University

*Starbucks is not affiliated with this study.

APPENDIX C

Follow-up Survey Reminder

Dear past or current student in the educational leadership program at VCU,

Reminder: Request for Help

This email is to thank you if you completed the survey for my dissertation research regarding classroom behavior management. If you have been unable to complete the survey for any reason, and are still willing to participate, clicking the link below and accepting the informed consent will allow you to begin the survey (insert link)

If you would like to participate, please respond by 11:00 p.m. on June 7, 2019. Access to the survey will expire after this time. You will have fifteen minutes to complete this survey once clicking the link.

If you have questions at any time about the survey or procedures, you may contact Kierstyn Johnson-Wigfall by email at johnsonkk2@mymail.vcu.edu.

Thanks again for your assistance,

Kierstyn Johnson-Wigfall
Virginia Commonwealth University

APPENDIX D

Follow-up- Final Survey Reminder

Follow-up Final Survey Reminder

Dear past or current student in the educational leadership program at VCU,

Final Request for Help

This final email is to thank you if you completed the survey for my dissertation research regarding classroom behavior management. If you have been unable to complete the survey for any reason, and are still willing to participate, clicking the link below and accepting the informed consent will allow you to begin the survey (insert link). You will have fifteen minutes to complete this survey once clicking the link.

If you would like to participate, please respond by 11:00 p.m. on June 7, 2019. Access to the survey will expire after this time.

If you have questions at any time about the survey or procedures, you may contact Kierstyn Johnson-Wigfall by email at johnsonkk2@mymail.vcu.edu.

Thanks again for your assistance,

Kierstyn Johnson-Wigfall
Virginia Commonwealth University

APPENDIX E

Electronic Survey Links

Control survey without student picture: <https://vcuportal.questionpro.com/t/AMjV8ZeUcz>

Light skin black male: <https://vcuportal.questionpro.com/t/AMjV8ZeUc1>

Dark skin black male: <https://vcuportal.questionpro.com/t/AMjV8ZeUDY>

Medium skin black male: <https://vcuportal.questionpro.com/t/AMjV8ZeUc2>

White male: <https://vcuportal.questionpro.com/t/AMjV8ZeUc0>

Light skin female male: <https://vcuportal.questionpro.com/t/AMjV8ZeUdC>

Dark skin female male: <https://vcuportal.questionpro.com/t/AMjV8ZeUc8>

Medium skin female male: <https://vcuportal.questionpro.com/t/AMjV8ZeUdB>

White female: <https://vcuportal.questionpro.com/t/AMjV8ZeUc5>

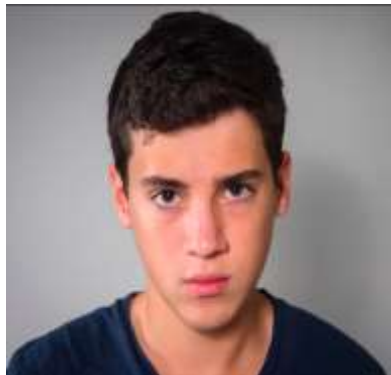
APPENDIX F

Original Student Photo

Female Student Photo



Male Student Photo



Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

APPENDIX G

Research Questions Statistical Significance Matrix for Student Skin Tone

	Scenario One By Harsh Discipline Decision	Scenario Two By Harsh Discipline Decision	Scenario Three By Harsh Discipline Decision and Student	Scenario Four By Harsh Discipline Decision
Light, medium, dark, white students and control group	No statistically significant differences	Statistically significant difference	No statistically significant differences	No statistically significant differences
Light, medium, dark and white students (no control group)	No statistically significant differences	No statistically significant differences	No statistically significant differences	No statistically significant differences
Light, medium, dark students and control group (no white students)	No statistically significant differences	Statistically significant difference	No statistically significant differences	No Statistically significant differences
Light, medium, dark students (no control group, no white students)	No statistically significant differences	No statistically significant differences	No statistically significant differences	No statistically significant differences
White students with control group	No statistically significant differences	No statistically significant differences	No statistically significant differences	No statistically significant differences
Students of color, white students and control group	No statistically significant differences	Statistically significant difference	No statistically significant differences	No statistically significant differences
Students of color and control group	No statistically significant differences	Statistically significant difference	No statistically significant differences	No statistically significant differences
Students of color and white students (no control group)	No statistically significant differences	No statistically significant differences	No statistically significant differences	Statistically significant difference
All students and control group	No statistically significant differences	Statistically significant difference	No statistically significant differences	No statistically significant differences
Student Gender	No statistically significant difference	No statistically significant differences	No statistically significant differences	No statistically significant difference
Intersection of Student Gender and Skin Tone (students of color, white and control group)	No statistically significant differences	No statistically significant differences	No statistically significant differences	No statistically significant differences
Intersection of Student Gender and Skin Tone (students of color, white and no control group)	No statistically significant differences	No statistically significant differences	No statistically significant differences	***Statistically significant and meaningful
Intersection of Student Gender and Skin Tone (students of color only)	No statistically significant differences	No statistically significant differences	***Statistically significant and meaningful	No statistically significant differences

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Research Questions Statistical Significance Matrix for Student Skin Tone and Educator Race

	Scenario One By Harshest Discipline Decision	Scenario Two By Harshest Discipline Decision	Scenario Three By Harshest Discipline Decision and Student	Scenario Four By Harshest Discipline Decision
Intersection of Educator Race and Student Skin Tone	Analysis could not be completed due to small cell size	Analysis could not be completed due to small cell size	Analysis could not be completed due to small cell size	Analysis could not be completed due to small cell size
Educator Race	No statistically significant differences	No statistically significant differences	No statistically significant differences	No statistically significant differences
Educator Race and Brown Skin Tone	No statistically significant difference	No statistically significant difference	No statistically significant difference	No statistically significant difference
Intersection of Educator Race and Student Gender	Statistically significant difference	No statistically significant differences	No statistically significant differences	Statistically significant difference
Intersection of Educator Race by Student Skin Tone and Gender	Analysis could not be completed due to small cell size	Analysis could not be completed due to small cell size	Analysis could not be completed due to small cell size	Analysis could not be completed due to small cell size

Running Header: Impact of Students' Skin Tone and Gender on Educators' Disciplinary Decisions

Research Questions Statistical Significance Matrix for Student Skin Tone and Educator Gender

	Scenario one By Harshes Discipline Decision	Scenario Two By Harshes Discipline Decision	Scenario Three By Harshes Discipline Decision and Student	Scenario Four By Harshes Discipline Decision
Educator Gender	No statistically significant differences	Statistically significant difference	No statistically significant differences	No statistically significant differences
Intersection of Educator Gender and Student Skin Tone	No statistically significant differences	***Statistically significant and meaningful	No statistically significant differences	No statistically significant differences
Intersection of Educator Gender and Student Gender	No statistically significant differences	No statistically significant differences	No statistically significant differences	No statistically significant differences
Intersection of Educator Gender by Student Skin Tone and Gender	Analysis could not be completed due to small cell size	Analysis could not be completed due to small cell size	Analysis could not be completed due to small cell size	Analysis could not be completed due to small cell size