Implicit Bias Awareness and Intervention Influence on In-service Classroom Teachers Promoting Equity in School Discipline: A Mixed Methods Study

James L. Pate, Ph. D.
Lantry L. Brockmeier, Ph. D.
John Lairsey, Ed. D.
Denise L. Hill, Ed. D
Valdosta State University
Valdosta Georgia

Nicholas R. Chastian, Ed. D
Colquitt County Schools
Moultrie, Ga.

Abstract

Schools and districts across the nation have long sought pragmatic solutions to resolve problems associated with discipline disproportionality. This study examined the topic of equity in school discipline from the perspective of classroom teachers. Using the Teacher Multicultural Attitudes Survey (TMAS) and the race Implicit Association Test (IAT), 60 in-service classroom teachers completed the quantitative portion of the study while nine teachers participated in the qualitative portion of the study. Results from the quantitative portion of the study were not statistically significant, yet several issues of practical significance were identified. Teacher repeated measure scores on the IAT indicated a slight change in preference from White to Black skin. In contrast, teacher pretest and posttest scores on the TMAS indicated teachers had less awareness of and sensitivity to multicultural issues in their classroom. Interview data, however, indicated teachers seeking to improve their efficacy were willing to discuss and address this problem.

Key Words: equity, school discipline proportionality, cultural awareness, implicit racial bias, teacher efficacy

1. Introduction

Discipline disproportionality is a problem plaguing many schools in the nation – especially the Southeastern United States (U.S. Department of Education, 2016a). The U.S. Department of Justice and Education defines discipline disproportionality as punishing a racial or ethnic subgroup of the population at a higher rate than would be expected (2014). The U.S. Department of Education's Office for Civil Rights releases data annually that describe and measure equity in school discipline in the United States (2016b). Losen, Hewitt, and Toldson (2014) reported the number of schools able to reduce or eliminate discipline disproportionality is few. The Civil Rights Data Collection (U.S. Department of Education, 2016b) reported African American students were suspended at nearly four times the rate of European American students in the United States during the 2013-14 school year. The U.S. Department of Education (2016a) claimed students who were removed from the learning environment for discipline reasons often face negative consequences possibly affecting them long after they become adults. Students who were suspended were more likely to drop out of school and never graduate, fell further behind academically, and were more likely to be suspended again. Losen et al. (2014) reported there is no systemic policy or guideline proven to reduce discipline disproportionality. Gregory, Skiba, and Noguera (2010) assert there is no current or historical research definitively helping schools reduce discipline disproportionality. Gregory et al. alleged the problem is getting worse.

2. Literature Review

Equity in school discipline (discipline disproportionality) is a problem schools in the United States have grappled with for decades (Children's Defense Fund, 1975; Gordon, Piana, &Keleher,2000; U.S. Department of Education, 2016a). Discipline disproportionality is the systemic over-representation of a particular subgroup or ethnicity of students when analyzing school, district, or state level discipline data (U.S. Department of Education, 2014). Although discipline disproportionality can be measured in several ways, one fact remains – African American students are three to four times more likely to receive an office discipline referral (ODR) than their European American peers (Brown & Steele, 2015; Freeman & Steidl, 2016; Krezmien, Leone, & Achilles, 2006; McIntosh, Ellwood, McCall, & Girvan, 2017; Okonofua & Eberhardt, 2015; U.S. Department of Education, 2016b).

Students receiving an ODR inherently lose time in the classroom. This loss of instructional time has a negative effect on student achievement and increases the risk a student will drop out of high school (Bleyaert, 2009; Fisher, 2011; Gass & Laughter, 2015; Gordon et al., 2000; Irby, 2013; Monahan, Vanderhei, Bechtold, & Cauffman, 2014; Nance, 2016; Noltemeyer, Ward, &McLoughlin, 2015; Shollenberger, 2015). More to the point, punitive discipline practice has been linked to higher rates of involvement with the juvenile justice system and increases the likelihood of being incarcerated as an adult (Nance, 2016; Nicholson-Crotty, Birchmeier, & Valentine, 2009; Noltemeyer et al., 2015; Shollenberger, 2015). Nicholson-Crotty et al. (2009) found African American students in Missouri were more likely to face punitive discipline measures than their European American peers, even after controlling for environmental differences like poverty and urban density. Skiba, Michael, Nardo, and Peterson, (2000) found after controlling for socioeconomic status and gender, African American males were more likely than their European American peers to receive an ODR – often for subjective reasons. Skiba et al. theorized this discrepancy may be due to systemic cultural biases occurring at the classroom level. Although this problem spans generations and has been studied extensively, pragmatic effective solutions remain elusive (Losen et al., 2014).

Beck and Muschkin (2012) compared behavior and achievement data for African American and European American students in North Carolina. They found as the discipline gap increased (e.g. more ODR's for Black students), so did the achievement gap. Beck and Muschkin coined the term cradle-to-prison pipeline and argued teacher perception of African American students, even at an early age, creates enduring consequences for affected students. More recently, Pearman II, Curran, Fisher, and Gardella (2019) utilized data from the Stanford Education Data Archive and the Civil Rights Data collection. Analogous to Beck and Muschkin, Pearman II et al. found the discipline gap and the achievement gap were highly correlated for African American students. As the discipline gap between African American and European American students increased, so did the achievement gap. Additionally, Pearman II et al. found the opposite was true. African American student achievement increased as disproportionate discipline practices decreased. The results of this study support Shollenberger's (2015) research. She found the achievement gap decreased significantly for African American students after controlling for suspension. Consequently, Shollenberger found any student (African American or European American) suspended over 10 days was more likely to be arrested or drop out of high school. Pearman II et al. suggest districts working to improve the discipline gap may get the unintended benefit of closing the achievement gap (or vice versa).

Similarly, Sullivan and Bal (2013) found African American students were more likely to be under identified for gifted services and over identified for special education services. The U.S. Department of Education (2016a) defines special education disproportionality in the same way they define discipline disproportionality. The U.S. Department of Education utilizes a risk ratio analysis comparing the rate one particular subgroup of students has of being identified for special education services (or receiving an ODR) compared to a comparison subgroup. To that end, Sullivan and Bal believe the two are highly correlated. They assert African American students receiving exclusionary discipline as a punitive measure inherently miss instruction, fall behind their peers academically, and subsequently endure lowered teacher expectations. Sullivan and Bal argue this cycle leads to more African American students being identified for special education services and a litany of other negative consequences (negative teacher perception, higher dropout rate, old for grade, etc.). Furthermore, Vanderhaar, Munoz, and Petrosko (2014) discovered African American students entrenched in this cycle are more likely to be referred to an alternative placement or school for disciplinary reasons.

In like fashion, Hughes and Kwok (2007) found African American students were more likely to have negative teacher relationships. They suggest negative teacher/student/parent relationships were largely responsible for the achievement gap between African American students and their peers. Indeed, Mortenson (2018) reported lower teacher expectations and negative teacher perceptions of African American students were highly correlated with the achievement gap.

Reducing discipline disproportionality and eliminating the achievement gap have both been prioritized for years with limited success (Losen et al., 2014). Blitz, Anderson, and Saastamoinen (2016) found teachers in their study were highly sympathetic to the impoverished lives many of the African American students in their classrooms face. Interestingly, the same teachers considered African American students more apathetic, aggressive, and disruptive than other students. When teachers were required to participate in cultural awareness professional development, many were insulted by the insinuation they were culturally deficient in some way (Blitz et al., 2016). Blitz et al. reported many of the teachers claimed to be colorblind while teaching – they did not consider race when reviewing student achievement and classroom discipline data. Although Sue et al. (2007) asserted colorblindness is considered a racial micro invalidation (racist ideology), Hartmann, Croll, Larson, Gerteis, and Manning (2017) claim the colorblind ideology has evolved for many.

Hartmann et al. believe individuals struggling with racial inequity and who claim to be colorblind are emotionally sensitive and should be treated with care when facing personal issues like race and culture. Consequently, forcing teachers to undergo cultural awareness training is unlikely to be an effective measure at reducing the discipline gap (Blitz et al., 2016).

Boneshefski and Runge (2014) think having teachers examine discipline data is a promising strategy to reduce discipline disproportionality. They contend by examining school level discipline data, teachers might become aware of their discipline tendencies and self-correct. Amin (2017) theorized a similar outcome for judges who examine the racial data of their courtroom decisions. To this end, helping individuals understand how implicit bias affects decision making, in a non-threatening manner, is a promising approach to reduce inequity in school discipline.

The psychological premise of implicit bias has been studied for decades (Greenwald &Banaji, 1995). Greenwald and Banaji believe when individuals are faced with making an automatic response to a given situation, implicit bias may cause them to make a different decision than they would if they had the time to carefully consider the situation. Further, Greenwald and Banaji assert many individuals are not aware of their implicit biases and the possible negative consequences their biases may have on others. This is particularly interesting when considering discipline disproportionality. Teachers are often faced with making split-second decisions in their classroom concerning student behavior and discipline.

Implicit racial attitudes have been identified in children as young as three (Qian et al., 2016; Setoh et al., 2017). Interestingly, Qian et al. found implicit racial attitudes seemed to evolve over time. More to the point, Setoh et al. (2017) reported children not yet aware racial bias was socially undesirable had similar levels of explicit and implicit racial attitudes (while the adults did not). These findings indicate implicit racial attitudes develop early in life and change over time. Glock and Klapproth (2017) found teachers in Germany held implicit racial biases against Turkish students (an ethnic minority) while Hannon, Defina, and Bruch (2013) found teachers preferred African American female students with lighter skin over African American female students with darker skin. Similarly, Okonofua and Eberhardt (2015) found teachers were more likely to hold lower regard for hypothetical students based on student name alone (e.g. – James vs. Jamaal).

Although the literature on implicit racial bias is growing, the mechanism to help individuals become aware of possible bias is questionable. Dovidio and Gaertner (2000) believe educated professionals are well aware of the negative connotations surrounding explicit racial attitudes. Smolkowski, Girvan, McIntosh, Nese, and Horner (2016) and McIntosh et al. (2017) purport teachers, in particular, are acutely aware of this phenomenon and will make every effort to appear racially non-biased in their classroom and community. Notwithstanding, Smolkowski et al. believe when teachers are unaware of their implicit bias and face split-second decisions in their classroom or school, the decision is likely tainted with bias. McIntosh et al. propose teachers may be able to reduce their bias if they are made aware of it and are given strategies to mitigate the negative effects bias may have on their discipline decisions.

Intervention designed by Devine, Forscher, Austin, and Cox (2012) was designed to make individuals aware of their bias and gives them strategies to reduce or overcome their bias. The intervention includes a series of vignettes describing situations where bias may be present and provides individuals strategies designed to reduce or eliminate bias in a systematic manner. The intervention was found effective at reducing bias as measured by the Race IAT in undergraduate psychology students (Devine et al., 2012). The Race IAT is a latency based cognitive assessment that measures implicit bias through association (black skin/pleasant image vs. white skin/unpleasant image or vice versa). Pepis (2017) conducted a similar study using the Devine et al. intervention with pre-service teachers. Although her study did not produce statistically significant results, participants were able to lower their bias as measured by the IAT over time.

3. Methodology

A sequential explanatory mixed methods design was used in this study. This design was employed to investigate if teacher awareness of implicit bias changed over time after being exposed to an intervention designed to reduce or eliminate bias and teacher perception of the process used to make them aware of their bias. Once the quantitative data were analyzed, interview questions were developed to better understand teacher perception of the study and explain the largely non-significant findings from the quantitative portion of the study. Creswell and Plano Clark (2017) state this design "can be used to explain the mechanisms through qualitative data that shed light on why the quantitative results occurred and how they might be explained" (p. 77). For the quantitative methods of data collection, volunteer in-service teachers participated in three online Qualtrics surveys over three weeks. The specific data collected and analyzed varied by the control and experimental group. Both the control group and experimental group took the Teacher Multicultural Attitudes Survey (TMAS) as a pretest and posttest measure.

The experimental group was asked to take the race Implicit Association Test (IAT) at time points 1, 2, and 3. The qualitative method of data collection included the interviews of nine teachers who participated in the quantitative portion of the study. Creswell and Plano Clark (2017) describe this variation of the explanatory sequential design as the "follow-up explanations variant" (p. 82). Five participants from the experimental group and five participants from the control group indicated they were willing to be interviewed for the qualitative portion of this study. One participant from the control group was unable to participate. A one-way analysis of covariance was used to answer research question 1, is there a significant difference between pretest and posttest scores on Ponterotto's (1995) Teacher Multicultural Attitude Survey (TMAS) by the control group and experimental group? A one-way analysis of variance was used to answer research question 2, is there a significant difference among participants who score low or high on the pretest TMAS on the final IAT score, and a within-subject repeated-measures analysis of variance was used to answer research question 3, is there a significant difference in participant repeated measure scores on the Implicit Association Test (IAT)?

The demographic data gathered from the 60 participants who completed the pretest and posttest TMAS indicated most were female (57), White (55), and taught at the elementary level (37). Participants ranged in age from $25 - 55+$, and reported varying years of experience. Before completing the Devine et al. (2012) intervention, 52 of the participants were concerned about equity in school discipline, 50 believed equity in school discipline affected the academic outcomes of students in their school, and 48 thought equity in school discipline affected the climate of their school. After completing the study, 45 participants were concerned with equity in school discipline, 48 believed equity in school discipline affected the academic outcomes of students in their school, and 44 thought equity in school discipline affected the climate of their school. Fifty-one participants indicated they thought participating in the study would influence their professional practice. Fisher's (2011) exact tests were conducted to compare the participants before the pretest and after the posttest TMAS by the control group and experimental group. The groups were compared on their level of concern on equity in school discipline, if equity in school discipline affected the academic outcomes of their students, and if equity in school discipline affected the climate of their school. Only one Fisher's exact test was significant. After taking the posttest TMAS, participants in the control group were significantly more concerned about the topic of equity in school discipline than the experimental group ($p = .001$, 95% CI [2.08, 117], $OR = 11.2$).

4. Results

The results of this study are reported examining the investigation if teacher awareness of implicit bias changed over time after being exposed to an intervention designed to reduce or eliminate bias and teacher perception of the process used to make them aware of their bias. Both quantitative and qualitative results are presented in this sequential explanatory mixed methods study.

4.1 Quantitative Findings

Research question 1 was designed to measure if teacher awareness of and sensitivity to multicultural issues in their classroom changed over time by the control group and experimental group. All participants took the Teacher Multicultural Attitude Survey (TMAS). Experimental group participants were also asked to complete the Race Implicit Association Test (IAT). The TMAS is a 20 item Likert-type instrument designed specifically for in-service classroom teachers. Scores can range from $20 - 100$. A score of 20 would indicate a teacher has little or no appreciation and awareness of multicultural teaching issues in their classroom while a score of 100 would indicate a teacher is highly appreciative and aware of multicultural teaching issues in their classroom. A one-way analysis of covariance was used to answer research question one. The overall mean pretest score for all 60 participants was $M = 82$ ($SD = 7.32$), higher than the overall mean posttest score for all 60 participants ($M = 81$, $SD = 8.19$). Control group participants ($n = 34$) had an overall mean pretest score of $M = 83.88$ ($SD = 6.29$) and overall mean posttest score of $M = 82.74$ ($SD = 6.93$). Experimental group participants ($n = 26$) had an overall mean pretest score of $M = 79.54$ ($SD = 7.94$) and overall mean posttest score of $M = 78.73$ ($SD = 9.23$). Interestingly, TMAS scores for all the participants, control group participants, and experimental group participants trended down from pretest to posttest. This indicates participants had less appreciation and awareness of multicultural teaching issues in their classroom after participating in the study. A one-way analysis of covariance (ANCOVA) was then conducted to determine if there was a significant difference in posttest TMAS scores by the control and experimental group after controlling for pretest TMAS scores. After adjusting for pretest TMAS scores, there was not a significant difference in posttest TMAS scores between the control and experimental group (F)(1, 57) $= 0.07$, $p = .80$).

Research question 2 investigated if participants who scored low or high on the pretest TMAS had significantly different IAT scores at time point 3 (IAT3). IAT scores are bound at the ±2 level.

Race IAT scores with a positive value indicate an individual has a preference for white skin, and Race IAT scores with a negative value indicate a preference for black skin, respectively. This question was limited to the seven experimental group participants who completed IAT assessments at time point 1, 2, and 3. Demographic characteristics of the seven participants who met this requirement indicated six teachers were female, White, taught at the elementary level, and had over 15 years of experience. The teachers in this group included two with a bachelor's degree, two with a master's degree, and three with an education specialist's degree. The low group ($n = 4$) pretest TMAS mean was $M = 73$ and high group ($n = 3$) pretest TMAS mean was $M = 85.33$. The low group mean IAT3 score was $M = 0.14$ and high group mean IAT3 score was $M = -0.29$. This signifies teachers who scored low on the pretest TMAS had little to no preference for white skin at time point three, and teachers who scored high on the pretest TMAS had a slight preference for black skin at time point three. The one-way ANOVA, however, indicated there was not a significant difference on mean IAT3 scores among participants who scored low or high on the pretest TMAS $F (1, 5)$, $p = .29$.

Research question 3 examined if participant IAT scores changed over time. A within-subject repeated-measures ANOVA was computed to investigate whether there was a significant difference in participant repeated measure IAT scores. Only the seven experimental group participants who completed all 3 IAT assessments were included in this analysis. The overall mean IAT1 score was $M = 0.42$, the overall mean IAT2 score was $M = 0.09$, and the overall mean IAT3 score was $M = -0.04$. The within-subject repeated-measures ANOVA indicated there was no significant difference in participant repeated measure IAT scores $F (2, 12) = 2.23$, $p = .15$).

Although no significant difference was found in repeated measure IAT scores, it is worth noting scores did trend down over time. At time point 1, the mean IAT score was 0.42, indicating most participants had a moderate preference for white skin. At time point 2, the mean IAT score dropped to 0.09, indicating individuals had little to no preference for individuals with white skin. Remarkably, at time point 3, individuals had a mean IAT score of -0.04, indicating individuals had little to no preference for individuals with *black* skin. This suggested although the change in scores was not statistically significant, the underlying premise of implicit bias changing over time *did* occur. As measured by the IAT, participants started the study with a moderate pro-white skin bias and ended the study with little to no pro-black skin bias. This finding would suggest making concerned individuals aware of implicit bias and giving them strategies to reduce or eliminate their bias, was successful. Devine et al. (2012) agreed this change in preference is significant and lends credence to her belief individuals must be concerned about implicit bias and motivated to change before they can reduce or eliminate their bias. Interestingly, participant *appreciation and awareness* about multicultural teaching issues in the classroom, as measured by TMAS mean scores, decreased for both the control group and experimental group participants in the study.

4.2 Qualitative Findings

Both control and experimental group participants were included in the interview portion of this study. Interviewees were able to view the questions and review the results of the quantitative portion of the study before being interviewed. The questions centered on teacher perception of the surveys and the Devine et al. (2012) intervention, the differential dropout rate between the control and experimental groups and the change in preference from white skin to black skin as measured by the Race IAT. In addition, the questions examined the significance of the study concerning equity in school discipline, teacher effect on students, personal reflections concerning implicit bias, the role of K-12 leadership in addressing implicit bias, and thoughts on if training in this area would affect the teaching and learning at their school.

Interview responses on what was most important to control group and experimental group participants revealed two distinct themes. Experimental group participants all reported taking the Race IAT was the one aspect of the study they remembered most. This reaction was somewhat expected according to research indicating the results from the Race IAT may cause participants to experience powerful emotions of distress, discomfort, disregard, disbelief, acceptance, or react defensively (Clark & Zygmunt, 2014; Howell, Gaither, & Ratliff, 2014). Several respondents admitted the IAT surprised them, explaining they were not expecting the images and words to switch in the middle of the IAT. Olivia confessed when the words and pictures "switched", it took a minute for her mind to adjust. Experimental group participants also talked about the emotions they felt after taking the IAT. Amelia stated, "I felt bad about myself thinking that I might have some biases." All respondents indicated the IAT made them more aware of implicit bias. Ava remarked "I'd like to take it (IAT) again."

Control group participants, conversely, hardly mentioned the IAT when reflecting on what they remembered most about the surveys. Participants from the control group thought the Devine et al. (2012) intervention was effective at increasing their awareness of implicit bias. The intervention did seem to increase participant's awareness of bias and the overall concern individuals had on how these biases may negatively affect their students (or society at large).

James mentioned his overall "self-awareness" increased and Mary expressed one of her "biggest fears (as a teacher) is that I would have some… hidden racism or implicit bias and not realize it." Overall, the respondents in the control group expressed feelings of curiosity and a desire to improve as driving factors to participate in the study and considered the intervention and experience of participating professionally beneficial.

Experimental group participants were divided on why they thought the control group mean TMAS scores were higher than the experimental group mean scores. Charlotte and Ava were not surprised the control group mean TMAS scores were higher than their group. Charlotte figured most professionals can "control what they are thinking" and will usually say the right thing to appear professional or non-racist. Although she did not explicitly mention the IAT or the anonymous nature of the surveys, her response indicated she thought the scores were lower for her group because of the IAT. She considered herself colorblind, stating she does not see "color and races" she sees "children" in her classroom. Ava also thought the scores were different because of the IAT. She mentioned the IAT did not affect her as much as other teachers, because as an ESOL (English to Speakers of Other Languages) teacher, she considered herself culturally competent. Amelia, Olivia, and Aurora, however, were perplexed by the difference in group scores. They figured TMAS scores would have increased for the experimental group *because* of the IAT. Amelia put it plainly: "I would have thought the experimental group (TMAS score) would have gone up and made the overall score higher for the experimental group." Amelia considered the groups may have just been made up differently and one group was more culturally sensitive than the other. Aurora thought about it for a moment, and after realizing the control group did not take the IAT, believed control group participants were not as honest on the TMAS as experimental group participants and "were basically assessing themselves."

Nearly all control group participants explicitly stated they thought the difference in group mean TMAS scores was due to the IAT. Mary thought the IAT probably affected experimental group participants more than usual due to the ongoing civil unrest the United States was experiencing during the summer of 2020. Jennifer thought the control group participants were not as honest as experimental group participants and answered TMAS questions how they felt they *should* answer rather than answer honestly. John also figured the IAT had something to do with the difference in scores and predicted experimental group participants were probably "more aware" of their personal biases. Interestingly, both groups mean TMAS scores dropped slightly. TMAS scores are considered to measure teachers' explicit attitudes about multicultural issues in the classroom. This indicates participants rated their explicit racial attitudes less favorably after participating in the study. Devine et al. (2012) theorized the intervention would cause participants to become more culturally aware and care more about the unintended consequences implicit bias causes for affected groups. This data suggests as awareness and knowledge is gained by utilizing the intervention, individuals realize they may not be as culturally neutral as they thought.

Considerably more experimental group participants dropped out of the study either partly or all together than did control group participants. Both groups were divided on if they considered the dropout rate reasonable or explainable. This data strongly correlates with the findings of both Clark and Zygmunt (2014) and Howell et al. (2014). They found if individuals were not prepared to receive their IAT results or if IAT results were extremely different than their self-reported levels of explicit bias, they would most likely disengage from the process or become defensive and more entrenched in their previous beliefs (Clark & Zygmunt, 2014; Howell et al., 2014).

A few experimental and control group participants did not think the dropout rate difference was reasonable or explainable. They did not consider dropping out of the study and were curious why some participants would finish the surveys but not the IAT. Amelia did consider the IAT "frustrating" when her scores were different than what she wanted them to be and Olivia thought some participants did not like taking the IAT multiple times. James thought most teachers should be professionally willing to grapple with issues like implicit bias. In a recent study, however, Starck, Riddle, Sinclair, and Warikoo (2020) found teachers were like other professionals in both implicit and explicit bias measures. Nearly all professionals in the two national data sets they utilized in their study held some degree of pro-white implicit and explicit racial attitudes (Starck et al., 2020). These findings, Starck et al. suggest, should encourage schools and teachers to investigate ways to help teachers reduce or eliminate either bias in general or the effects of bias on students in particular (2020).

Nearly all participants from both the control group and experimental group considered the experience of participating in the study personally significant and affected their views on the topic of equity in school discipline. Specifically, they attributed the change in preference from white skin to black skin due to teacher effort. More precisely, they believe the teachers who completed the IAT three times were trying to drop their score. The rationale, however, varied among the respondents. Experimental group participants Ava and Amelia considered it a professional responsibility to lower their IAT scores.

Amelia thought "everybody (teachers) is trying to be more racially sensitive" and Ava thought it part of a teachers "job to learn other people's culture." Olivia and Aurora considered it a cognitive challenge to lower their scores. Olivia stated the second and third time she took the IAT, "I was prepared for the changes (key switches for black/white and pleasant/unpleasant)." Aurora thought the IAT made her "sit there and think", she knew what the assessment was trying to measure and how it was measuring it. Aurora thought teachers might be given IAT assessments other than race in the future then hesitated: "but, then, you want them to be more conscious of it (racial implicit bias)." Control group participants James and Jennifer thought teachers in the experimental group intentionally tried to lower their scores. James thought teachers stopped seeing color and "focused on the keys" and Jennifer posited the inherent experience of taking the IAT and surveys could "push you (to change)." She thought as individuals became aware of their bias and acknowledge it, true change would manifest. Control group participants Mary and John did not weigh in on this change. Mary studied the Race IAT to understand how it measures implicit bias and questioned the validity of the instrument.

Charlotte, from the experimental group, provided the only disconfirming answer when compared to all other control and experimental group participants on the topic of multicultural teaching issues effect on students. She believes too much emphasis has been placed on the topic of race in our society. She sees "a lot of problems that has [*sic*] absolutely nothing to do with race." It is worth noting the intense, year-long program Gregory et al. (2016) found effective at reducing the discipline gap and closing the achievement gap does not "explicitly focus on raising teacher consciousness about implicit bias or institutional racism… rather, it focuses on skills in effectively interacting with any student" (p. 186). Experimental group participants Ava, Amelia, Olivia, and Aurora thought students were affected in several ways by the issues of race, culture, and bias (Blitz et al., 2016). Ava and Oliva believe students should be taught to be more tolerant of their peers and Amelia suggested many of the issues are taught and learned in environments outside the classroom. All four control group participants thought the issues of race, culture, and bias had a direct negative impact on their students. James mentioned the gang culture prevalent in a neighborhood near his elementary school while Mary and Jennifer believe lower teacher expectations and subsequent academic disengagement many African American students face create a culture of failure. John, who teaches at the high school level, thought the overwhelming amount of news coverage over the summer of 2020 would cause cultural issues at his school. The Black Lives Matter movement and the never-ending news cycle of White police officers shooting unarmed Black men, John continued, was bound to have an impact – especially when African American students look around their school and all they see "are a bunch of White teachers."

Almost all participants were able to recall and share a personal story concerning implicit bias in their school or classroom. Ava, Amelia, and Olivia's stories were deeply personal and affect them still. Ava shared that her relationship with an African American student blossomed after she slowed down and listened to him. She confessed that memory influences her classroom practice daily. Amelia and Olivia still worry about how their decisions may have had long term consequences or negative effects for the students they described. Control group participant Jennifer spoke about her first year as a teacher in a school with predominantly African American students. James, Mary, and John all spoke about the importance of building relationships with the students in their classrooms.

Several respondents from the experimental group and all control group participants thought K-12 leaders should take the time to build honest and open professional relationships with their teachers so meaningful conversations could occur. Charlotte and Ava from the experimental group were not asked this question. Amelia thought before teachers are comfortable thinking about implicit bias, they need to be comfortable talking about it with their building administrator. Aurora agreed, adding teachers and leaders need to have conversations about the topic. As an example, Aurora mentioned the civil unrest in the United States during the summer of 2020 stating, "everything going on in the nation, you know, you feel bad sometimes for being white." James put the onus on leaders to become cognizant of the effects implicit bias may have on the students at their school and model the correct behavior for teachers. He believes leaders need "to make it a point" and "make it a practice" for the teachers in their school. Mary thought clear expectations concerning behavior and discipline from school leaders would alleviate many issues stating "my God – don't sweat the small stuff!"

The teachers had varying opinions on how to implement effective training on the topic of implicit bias at their school. Some experimental group participants thought the training would need to be voluntary and pragmatic. Specifically, one suggested teachers would benefit from visiting the homes and communities of their African American students. Another asserted, however, mandatory teacher training is already overwhelming, but considered it might be helpful for teachers who "really want to improve." Contending implicit bias professional development should be a priority for all K-12 educators, one respondent indicated if teachers were able to attend training, they needed to be given adequate time to reflect on and practice what they learned.

In addition, control group participants were convinced training aimed at raising awareness on multicultural issues in the classroom would benefit the teaching and learning in their school and classroom. Specifically, as teachers become comfortable with and aware of implicit bias, effective analysis and problem solving could occur.

5. Limitations and Assumptions

The researcher has worked at the elementary, middle, and high school level as a teacher. All of the teachers in the study worked at Title I schools with high rates of poverty and ethnically diverse student populations. It should be noted teacher efficacy or perceived efficacy were not examined in this study.

The sample size of participants in the quantitative portion of this study were lower than expected and greatly limit the generalizability of the results. In addition, teachers interested in the topic of equity in school discipline were recruited to participate, further limiting the generalizability of the study. It should be mentioned the civil unrest surrounding the murder of George Floyd and subsequent racial riots in cities across the nation occurred during the quantitative data collection portion of this study. Furthermore, teachers in the study were not able to complete the 2019-2020 school year due to the novel coronavirus and subsequent closing of all public schools in the state.

The quantitative portion of the study ended in June of 2020 while the qualitative portion of the study, utilizing participants from the quantitative portion of the study, did not commence until August of 2020. This gap in data collection is another limitation of the study. Although typical for studies utilizing the explanatory sequential design, the timing could have been shortened so respondent memory of the surveys and IAT were clear.

6. Suggestions for Future Research

Interview data from both the control and experimental group illustrate the power of the IAT. Future research could include different types of the IAT (politics, gender, religion) with the Devine et al. (2012) intervention and Ponterotto's (1995) TMAS to explore the racial construct in a less threatening manner. Several participants also mentioned how becoming aware of implicit bias would improve their teacher efficacy. Therefore, it is suggested future iterations of this study examine student and teacher perception of efficacy as a pretest and posttest measure.

Similarly, future research could explore how K-12 leaders concerned about the topic of equity in school discipline utilizing a similar conceptual framework are effected. Finally, future research could replicate this study in other settings and locations with a larger teacher participant pool to investigate the generalizability of the results for teachers concerned about the topic of equity in school discipline.

7. Conclusion

This study includes some unique characteristics relatively new to the long-known problem of discipline disproportionality. By utilizing in-service, volunteer teachers concerned about the topic of equity in school discipline and providing teachers with an intervention designed to reduce or eliminate their bias, using an instrument specifically designed to measure teacher explicit racial attitudes, and the Race IAT for implicit racial bias, this study was able to build on previous research conducted by Devine et al. (2012) and Pepis (2017). By utilizing a control and experimental group, two dependent variables measured teacher change over time on the construct of prejudice. Additionally, nine teachers who participated in the quantitative portion of the study were interviewed to help understand the quantitative findings.

Research question 1 investigated if teacher appreciation and awareness of multicultural issues in the classroom as measured by the TMAS were significantly different by the control and experimental group. Research question 2 measured if experimental group participants who scored low or high on the pretest TMAS had significantly different IAT scores at time point 3. Research question 3 measured if teacher repeated measure IAT scores changed over time.

Although the literature on discipline disproportionality and equity in school discipline is substantial, the effect implicit bias may have on teacher decisions at the classroom level is yet emerging. This study has major implications for schools and districts seeking to improve in this area. Teachers seeking to improve their pedagogical efficacy seem willing to grapple with this socially sensitive topic when given the time and freedom to do so. If K-12 leaders address this topic through the lens of improved teacher efficacy, more teachers may be willing to examine racial disparities in their classroom and school. Although the quantitative findings of the study were not statistically significant, the change in teacher preference for the seven who completed the study for the experimental group from a moderate preference for white skin too little to no preference for black skin is intriguing. Coupled with the overall decrease in mean TMAS scores, these results were almost perplexing. The teachers who participated in this study were all on paid "leave" due to SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2).

They had the time and resources to participate. The differential dropout rate when comparing the control and experimental group and the extreme dropout rate of experimental group participants taking the IAT was a cautionary development.

The interview data made clear how powerful the race IAT is. Experimental group participants referenced the IAT repeatedly, while hardly mentioning the Devine et al. (2012) intervention. In contrast, control group participants remembered the Devine et al. intervention clearly and spoke on how it changed their awareness of implicit bias. Yet– the power of the IAT seemed to come at a cost. It is worth noting the experimental group mean TMAS pretest scores (*M* =79.54), taken before the initial IAT, were lower by any metric than the overall control group mean TMAS pretest scores (*M* = 83.88).

The interview data from respondents indicated their overall awareness and acknowledgment of implicit bias increased by participating. Charlotte notwithstanding, participants felt the process was a good use of their time professionally. Most striking, however, was the overwhelming desire to improve. Every teacher mentioned how important improving their teacher efficacy was to them professionally. Control group participant James summed it up:
Discipline disproportionality (leads to) lower reading, lower Lexile levels, lower achievement, higher poverty, higher teen pregnancy, higher mortality, the list goes on. And I could correlate every one of those findings with discipline. It is not about being scared. It's about making the change. So at some point the change has to be made, otherwise the issue will never go away.

Although this study failed to show a statistically significant drop in IAT scores over time, and teacher appreciation and awareness of multicultural issues in the classroom as measured by the TMAS dropped from pretest to posttest, it did give teachers concerned about the topic of equity in school a discipline a voice in this long raging debate. The qualitative data from this study indicate teachers seeking to improve will go to considerable effort to increase their efficacy. The fact teachers who completed the IAT three times were able to change from a moderate preference for white skin to little to no preference for black skin is remarkable. Even if teachers were trying to lower their scores for personal gratification, they were motivated to change their thought patterns. Finding a pragmatic solution to the problem of equity in school discipline is not likely to come in the form of a standalone, one-time intervention. More than likely, a combination of coaching teachers coupled with careful instruction on the topic of implicit bias will prove effective at reducing or eliminating inequity in school discipline. It will not be easy. If it were, the problem would have been solved long ago.

References

Amin, A. (2017). Implicit bias in the courtroom and the need for reform. *Geo. J. Legal Ethics, 30*, 575.

Beck, A. N., & Muschkin, C. G. (2012). The enduring impact of race: Understanding disparities in student disciplinary infractions and achievement. *Sociological Perspectives, 55*(4), 637-662. doi:http://dx.doi.org/10.1525/sop.2012.55.4.637

Bleyaert, B. (2009). *Disproportional high school suspension rates by race and ethnicity. Research brief*. Retrieved from http://library.valdosta.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED537954&site=ehost-live

Blitz, L. V., Anderson, E. M., & Saastamoinen, M. (2016). Assessing perceptions of culture and trauma in an elementary school: Informing a model for culturally responsive trauma-informed schools. *The Urban Review, 48*(4), 520-542. doi:http://dx.doi.org/10.1007/s11256-016-0366-9

Boneshefski, M. J., & Runge, T. J. (2014). Addressing disproportionate discipline practices within a school-wide positive behavioral interventions and supports framework: A practical guide for calculating and using disproportionality rates. *Journal of Positive Behavior Interventions, 16*(3), 149. doi: 10.1177/1098300713484064

Brown, K. E., & Steele, A. S. L. (2015). Racial discipline disproportionality in montessori and traditional public schools: A comparative study using the relative rate index. *Journal of Montessori Research, 1*(1), 14. doi:10.17161/jomr.v1i1.4941
Children's Defense Fund (1975). *School suspensions: Are they helping children?* Cambridge, MA. Retrieved from http://diglib.lib.utk.edu/cdf/data/0116_000050_000205/0116_000050_000205.pdf

Clark, P., & Zygmunt, E. (2014). A close encounter with personal bias: Pedagogical implications for teacher education. *The Journal of Negro Education, 83*(2), 147-161.

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research* (3rd ed.). Los Angeles, CA: Sage publications.

Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of experimental social psychology, 48*(6), 1267-1278. doi:10.1016/j.jesp.2012.06.003

Dovidio, J. F., & Gaertner, S. L. (2000). Aversive racism and selection decisions: 1989 and 1999. *Psychological Science (0956-7976), 11*(4), 315.

Fisher, J. D. (2011). *School composition and disproportionality in discipline: A school-level analysis of disciplinary actions against various offense types in the United States, 2007–2008* (Doctoral Dissertation). Retrieved from UMI. (3450545)

Freeman, K. J., &Steidl, C. R. (2016). Distribution, composition and exclusion: How school segregation impacts racist disciplinary patterns. *Race and Social Problems, 8*(2), 171-185. doi: http://dx.doi.org/10.1007/s12552-016-9174-9

Gass, K. M., & Laughter, J. C. (2015). "Can I make any difference?" Gang affiliation, the school-to-prison pipeline, and implications for teachers. *The Journal of Negro Education, 84*(3), 333-347.

Glock, S., & Klapproth, F. (2017). Bad boys, good girls? Implicit and explicit attitudes toward ethnic minority students among elementary and secondary school teachers. *Studies in Educational Evaluation, 53*, 77-86. doi:https://doi.org/10.1016/j.stueduc.2017.04.002

Gordon, R., Piana, L. D., & Keleher, T. (2000). *Facing the consequences: An examination of racial discrimination in U.S. public schools*. Retrieved from http://library.valdosta.edu:2048/login?url=http://search.proquest.com/docview/62339221?accountid=14800

Greenwald, A. G., Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*(1)*,* 4-27. doi:10.1037/0033295X.102.1.4

Gregory, A., Hafen, C. A., Ruzek, E., Mikami, A. Y., Allen, J. P., Pianta, R. C. (2016). Closing the racial discipline gap in classrooms by changing teacher practice. *School Psychology Review, 45*(2), 171-191.

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. doi: 10.3102/0013189X09357621

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. doi:http://dx.doi.org/10.1007/s12552-013-9104-z

Hartmann, D., Croll, P. R., Larson, R., Gerteis, J., & Manning, A. (2017). Colorblindness as identity: Key determinants, relations to ideology, and implications for attitudes about race and policy. *Sociological Perspectives, 60*(5), 866-888. doi:10.1177/0731121417719694

Howell, J. L., Gaither, S. E., & Ratliff, K. A. (2014). Caught in the middle. *Social Psychological and Personality Science, 6*(4), 373-381. doi:10.1177/1948550614561127

Hughes, J., & Kwok, O.-m. (2007). Influence of student-teacher and parent-teacher relationships on lower achieving readers' engagement and achievement in the primary grades. *Journal of Educational Psychology, 99*(1), 39-51. doi:10.1037/0022-0663.99.1.39

Irby, D. J. (2013). Net-deepening of school discipline. *The Urban Review, 45*(2), 197-219. doi:http://dx.doi.org/10.1007/s11256-012-0217-2

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

Losen, D., Hewitt, D., &Toldson, I. (2014). Eliminating excessive and unfair exclusionary discipline in schools: Policy recommendations for reducing disparities. *Discipline Disparities: A Research-to-Practice Collaborative*.

McIntosh, K., Ellwood, K., McCall, L., & Girvan, E. J. (2017). Using discipline data to enhance equity in school discipline. *Intervention in School and Clinic, 53*(3), 146-152. doi:10.1177/1053451217702130

Monahan, K. C., Vanderhei, S., Bechtold, J., & Cauffman, E. (2014). From the school yard to the squad car: School discipline, truancy, and arrest. *Journal of Youth and Adolescence, 43*(7), 1110-1122. doi:http://dx.doi.org/10.1007/s10964-014-0103-1

Mortenson, B. C. (2018). *The role of teacher implicit bias in the racial achievement gap* (Master's thesis). Retrieved from ProQuest. (10791341)

Nance, J. P. (2016). Dismantling the school-to-prison pipeline: Tools for change. *Arizona State Law Journal, 48*(2), 313-372.

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

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

Okonofua, J. A., & Eberhardt, J. L. (2015). Two strikes: Race and the disciplining of young students. *Psychol Sci, 26*(5), 617-624. doi:10.1177/0956797615570365

Pearman II, F. A., Curran, C. F., Fisher, B., & Gardella, J. (2019). Are achivement gaps related to discipline gaps? Evidence from national data. *AERA Open*. https://doi.org/10.1177/2332858419875440

Pepis, T. (2017). *Implicit bias in pre-service teachers: A mixed methods approach* (Doctoral Dissertation). Retrieved from http://scholarworks.umass.edu/dissertations_2. (959)

Ponterotto, J. G. (1995). Teacher multicultural attitude survey. New York, NY; Fordham University.

Qian, M. K., Heyman, G. D., Quinn, P. C., Messi, F. A., Fu, G., & Lee, K. (2016). Implicit racial biases in preschool children and adults from Asia and Africa. *Child Development, 87*(1), 285-296. doi:10.1111/cdev.12442

Setoh, P., Lee, K. J., Zhang, L., Qian, M. K., Quinn, P. C., Heyman, G. D., & Lee, K. (2017). Racial categorization predicts implicit racial bias in preschool children. *Child Development*, 1-18. doi:10.1111/cdev.12851

Shollenberger, T. L. (2015). Racial disparities in school suspension and subsequent outcomes: Evidence from the National Longitudinal Survey of Youth. *Closing the school discipline gap: Equitable remedies for excessive exclusion*, 31-44.

Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. (2000). *The color of discipline: sources of racial and gender disproportionality in school punishment*. Policy Research Report. Retrieved from http://library.valdosta.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED468512&site=eds-live&scope=site

Smolkowski, K., Girvan, E. J., McIntosh, K., Nese, R. N. T., & Horner, R. H. (2016). Vulnerable decision points for disproportionate office discipline referrals: Comparisons of discipline for African American and White elementary school students. *Behavioral Disorders, 41*(4), 178-195.

Starck, J. G., Riddle, T., Sinclair, S., & Warikoo, N. (2020). Teachers are people too: Examining the racial bias of teachers compared to other American adults. *Educational Researcher, 49*(4), 273-284. doi: 10.3102/0013189X20912758

Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A. M. B., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: Implications for clinical practice. *American Psychologist, 62*(4), 271-286. doi:10.1037/0003-066X.62.4.271

Sullivan, A. L., & Bal, A. (2013). Disproportionality in special education: Effects of individual and school variables on disability risk. *Exceptional Children, 79*(4), 475-494.

U.S. Department of Education. (2016a). *Office for Civil Rights. 2013-2014 Civil Rights Data Collection: A first look*. Retrieved fromhttps://www2.ed.gov/about/offices/list/ocr/docs/2013-14-first-look.pdf

U.S. Department of Education (2016b). Office of Special Education and Rehabilitation Services. *Racial and ethnic disparities in special education: A multi-year disproportionality analysis by state, analysis category, and race/ethnicity.* Retrieved from https://www2.ed.gov/programs/osepidea/618-data/LEA-racial-ethnic-disparities-tables/disproportionality-analysis-by-state-analysis-category.pdf

U.S. Department of Justice, Civil Rights Division and the U.S. Department of Education, Office of Civil Rights (2014). *Dear colleague letter: Nondiscriminatory administration of school discipline*. Retrieved from http://www.justice.gov/crt/about/edu/documents/dcl.pdf.

Vanderhaar, J., Munoz, M., &Petrosko, J. (2014). Reconsidering the alternatives: The relationship between suspension, disciplinary alternative school placement, subsequent juvenile detention, and the salience of race. *Journal of Applied Research on Children: Informing Policy for Children at Risk, 5*(2), 14.