

Teacher Education Students' Implicit Racial Attitudes and Interpersonal Attribution of Racialized Student Behavior

Nicole L. Lorenzetti
City University of New York
nlorenzetti@ccny.cuny.edu

Helen Johnson
City University of New York
hjohnson@qc.cuny.edu

Abstract

Teachers have been shown to hold lower behavioral expectations for Black students than for their White peers, and the mechanism underlying this may be teachers' implicit attitudes about their Black students based on causal attributions. This study examined this connection, predicting that teacher education students (TES) who scored higher on the racial implicit bias test would attribute internal causality and controllability to explain challenging behaviors in the classroom more frequently for Black students than for White students. 233 teacher education students completed the racial bias section of the Implicit Assessment Test and a set of questions assessing causal attribution based on four vignettes depicting student misbehaviors in a classroom setting. We found that regardless of implicit bias, TES were more likely to believe that Black students had an internal locus of causality and controllability than their White counterparts when presented with similar instances of challenging behavior. These results support the need for teacher preparation programs to address how these internal beliefs of teacher education students affect what they learn about managing their expectations around students' behavioral regulation and to what they attribute these behaviors.

Introduction

There are persistent racial disparities in the United States public school system in terms of both academics and discipline despite teacher education programs' claims of focusing on culturally responsive pedagogy (Gregory et al., 2010; Lacoé & Steinberg, 2019; Okonofua & Eberhardt, 2015; Skiba et al., 2011). While teacher education students' racial attitudes may be addressed explicitly in their preparation programs more often than in the past, there are still underlying beliefs about Black students that are not easily identified or openly addressed, as evidenced by the abundance of work on implicit racial attitudes (e.g., Okonofua & Eberhardt, 2015; Starck et al., 2020). Evidence supporting the construct of implicit racial attitudes generally, and teacher education students' attitudes specifically, highlights the need for teacher education programs to address these cognitive processes that are manifested in classroom practices, including through the hidden curriculum.

Attribution theory offers a lens for examining key dimensions underlying teacher racial beliefs, which ultimately impact teacher attitudes toward students. By juxtaposing implicit racial attitudes and attribution theory constructs using vignettes, the current study seeks to illuminate

factors – specifically implicit racial bias – that guide teacher education students’ (TES) approaches to deciding how to handle common types of misbehavior in the classroom by student race.

Teacher Attitudes

Teachers’ racial attitudes toward students in the classroom begin before they enter teacher education programs through their own racial development, as well as through their experiences interacting with the world (Starck et al., 2020). These racial attitudes are then cultivated through their teacher education programs, where they are often working within frameworks that claim to be culturally responsive (Little & Welsh, 2022). The development of racial attitudes toward students in teacher education programs is of particular value, as this is an opportune time to address how beliefs about and attitudes toward students shape how TES make both low-stakes and high-stakes decisions about students in the classroom on a daily basis.

Public school teachers in the United States are predominantly White, female, and hold at least a Bachelor’s degree (National Center for Education Statistics, 2023). According to Frankenberg (2012), these teachers’ racial attitudes are reflective of their extended ethnic group, which is explicitly supportive of educational integration. However, Frankenberg (2012) found that many White teachers attend homogenous colleges and universities, then return close to their equally homogenous home communities to teach (Frankenberg, 2012), particularly if their goal during their time in teacher preparation is to return to these communities as opposed to being open to teaching in a diverse environment.

Based on this lack of exposure to diversity, Frankenberg (2012) sought to examine if these teacher beliefs could be predicted by the teacher’s own race. The researcher used a national survey of 1,002 public school teachers from 48 states and found that teacher race significantly predicted their racial attitudes (Frankenberg, 2012). Individual teacher race was a significant predictor of race awareness, in that White teachers were found to espouse a more race-neutral attitude than teachers of color, and more homogenous student populations of schools predicted a race-neutral attitude in individual teachers in those schools: “in other words, in comparison to teachers in racially isolated White schools, teaching in racially isolated non-White schools predicted higher values of race-conscious attitude” (Frankenberg, 2012, p. 463).

Teacher beliefs, race-neutral or color conscious, about diverse student bodies affect teacher expectations of students, and ultimately impact how teachers interact with their students in the classroom. Teachers who anticipate behavioral issues and disordered learning environments in diverse urban schools may be inclined to implement control-focused rather than learner-centered classroom practices (Little & Welsh, 2022). Multiple studies “have found that teachers display negative attitudes toward, communicate low expectations to, and express, both verbally and nonverbally, negative and disapproving messages toward visible racial/ethnic children” (Carter & Goodwin, 1994, p. 307). These negative attitudes of teachers translate into behaviors toward students in the classroom and affect student outcomes, which can be seen through things such as tracking and other verbal and nonverbal interactions that shape decisions teachers make in the classroom (Carter & Goodwin, 1994; Papageorge et al., 2020).

Research shows that beliefs and attitudes become part of one's cognitive structure and then drive behavior, which may explain teachers' attitudes toward their students that affect their classroom interactions, as well as students' views of themselves as learners (Yang & Montgomery, 2013). Teachers bring their own worldviews and backgrounds into the classroom and see the world through their own racial identities (Castro, 2010). Racial identity is a psychological orientation toward one's own racial group, and it is not necessarily how one feels about membership, but how one psychologically views oneself within the group (Carter & Goodwin, 1994). Racial identity is often seen as irrelevant to members of the dominant race (White), who have rarely had to critically examine their own race, and it is an emotionally charged topic, making it a hot button issue (Carter & Goodwin, 1994). Lacking that critical examination, White teachers generally avoid, deny, or ignore racial issues in education (Carter & Goodwin, 1994; Gay & Kirkland, 2003), even though those issues negatively affect students of color in multiple ways, including through classroom rules, the curriculum, and larger school structures and systems (Carter & Goodwin, 1994, p. 292).

These teacher attitudes and beliefs inform how teachers interact with students in the classroom, and since White teachers still make up most educators in the United States (National Center for Education Statistics, 2023), these beliefs can trickle through their work, through both implicit and explicit bias, in schools that are racially diverse and in those whose composition of students are primarily students of color. Explicit bias is not always the biggest effector of expectations; racial attitudes also exist on an unconscious, or implicit, level and may have a large effect on how the teacher interacts with students (Cooper & Allen, 1998). Where do these expectations and attitudes come from in the cognitive processing of teachers? The answer may be in attribution theory.

Attribution Theory

The crux of attribution theory is that humans seek to make meaning out of information they receive from others, and they do that by attributing causes to behavioral outcomes to understand behavior. That understanding then prompts action from the attributor and allows them to predict future behavior (Hunter & Barker, 1987; Weiner, 2000). Weiner (2000) distinguishes between an *intrapersonal* theory of attribution, which is the attempt to understand oneself and the environment to explain causes of outcomes and drive future behavior, and an *interpersonal* theory of attribution, which one seeks to understand the causes of *others'* behaviors and environments in order to explain the actions and outcomes of the other. The focus of this study will be on the interpersonal theory of attribution, specifically in a school situation, where teachers seek to understand why students succeed or fail, and this gives rise to the application of attribution theory in the classroom (Weiner, 1979).

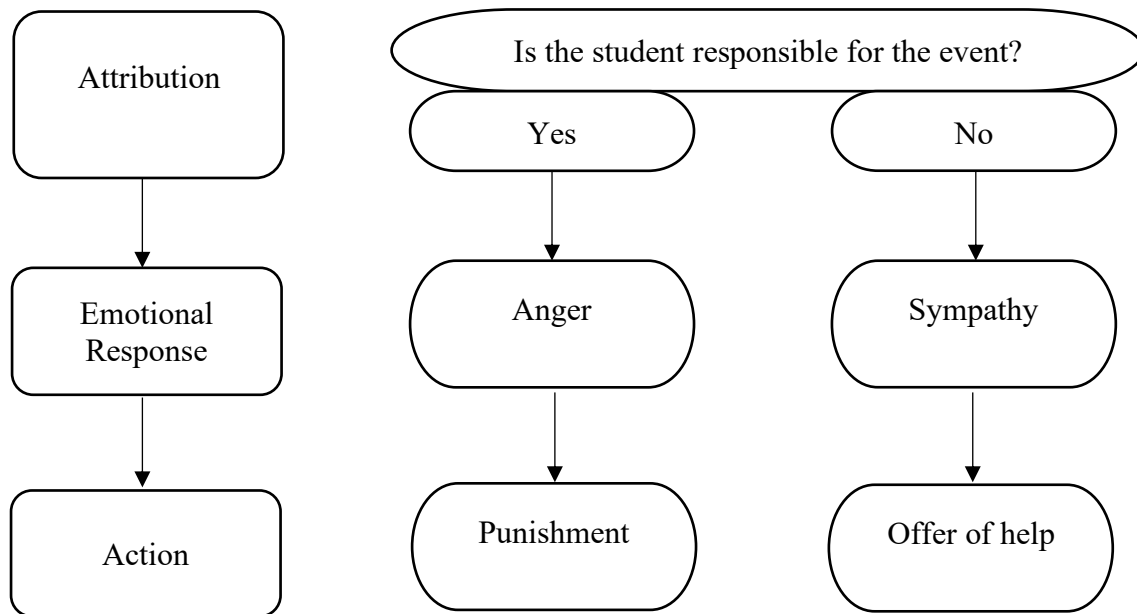
The key to attribution theory is that causation is in the eye of the beholder; humans *perceive* causes for the behavior of both themselves and others regardless of what the reality actually is (Hunter & Barker, 1987). This means that in interpersonal attribution, there may be a difference in perception of causes between the actor and the observer. In the classroom, this difference in perceptions may lead to conflicts between teachers and students (Hunter & Barker, 1987).

Attribution theory focuses on three dimensions: locus of causality, stability, and controllability. Locus of causality, often shortened to simply “locus,” is where the action begins, either internal or external to the person (Weiner, 1979, 1985, 2000, 2007, 2012). Stability is whether or not a cause can change; actors ask themselves whether or not what they can expect in the future is the same as what happened in the past (Hunter & Barker, 1987; Weiner, 2000). Stability can be either stable or unstable. Stable causes remain consistent over time, whereas unstable causes may or may not be present at any given time. Controllability is the belief in whether a person can willfully direct the outcome of an event (Weiner, 1979, 1985, 2000, 2007, 2012); Weiner (2000) refers to this as “volitional alteration” (p. 20). Controllability is also seen as moral responsibility, meaning that “individuals ‘ought’ to try hard, and they tend to be rewarded or punished to the extent that they exercise this responsibility” (Graham, 1988, p. 12).

If a teacher views a student as being in control of his or her behavior and a negative event occurs, the teacher may view the student as actively responsible for the event (controllable) as opposed to being a passive object of the event (uncontrollable) (Beckman, 1970; Graham, 1988; Hunter & Barker, 1987; Weiner, 2000). An example of a controllable trait is effort; an example of an uncontrollable trait is ability (Weiner, 1979, 1985, 2000).

Figure 1

Teacher reactions to interpersonal attributions toward student classroom behavior



The teacher’s reaction to the student’s behavior is then a result of attribution based on these three dimensions. If the teacher views the student as responsible for the event (internal locus of causality or belief in controllability), the teacher is likely to respond to the student with anger (Graham, 1988; Hunter & Barker, 1987; Weiner, 2000). Conversely, if the teacher views the student as having an external locus of causality or as lacking controllability, the teacher is likely

to respond to the student with sympathy (Graham, 1988; Hunter & Barker, 1987; Weiner, 2000). Anger often leads toward punishment, as opposed to sympathy, which may lead toward pro-social feedback such as offers of help (Weiner, 2000, 2007). Figure 1 presents this system.

Causal Attributions

According to attribution theory, causality beliefs guide current emotional states as well as future behavior (Weiner, 2000, 2007, 2012). In interpersonal attribution, anger and sympathy are the two dichotomous emotions that appear toward others, whereas happiness and shame are the two primary dichotomous emotions that are elicited in intrapersonal attribution (Graham, 1988; Weiner, 2000, 2012). Teachers' interpersonal attributions about students' behaviors affect students' intrapersonal theories of attribution (Beckman, 1970; Hunter & Barker, 1987). For example, if a teacher becomes angry at a student for failing a test because the teacher views the student as having high ability (an uncontrollable, internal, and stable trait), this may communicate to the student that he or she is responsible for the failure, and it is a personal failure, and this may induce either guilt or learned helplessness (Graham, 1988; Weiner, 2000).

This information on interpersonal attribution, when communicated intentionally or unintentionally to students, shapes their intrapersonal attribution belief system (Graham, 1988). For example, if a teacher expresses sympathy toward a student because she believes the cause of a student's behavior is uncontrollable, the teacher may be sending a cue to a student not just that the teacher believes the behavior in the student is uncontrollable, but also that the student can't be expected to do any better (Graham, 1988; Beckman, 1970; Weiner, 2012). This then shifts the intrapersonal belief system through social transmission. It is vital that teachers be aware of their interpersonal attributions and the effect they have on classroom emotions and student intrapersonal attributions. Graham (1988) has found evidence that children are better able to predict future success when they attribute causes to stable ones, and that students expect more blame from teachers when they view a cause as controllable.

Research has shown that teachers' views of student behavior, such as academic and behavioral attributions, affect how they treat students in the classroom (Chang & Demyan, 2007; Chang & Sue, 2003). Weiner (2000) describes punishment of students: "It is indeed the case that teachers primarily punish for utilitarian reasons [which focuses on future change]. However, when the cause is stable and controllable (i.e., the student is always lazy), then punishment is more directed toward justice, or giving 'what is deserved'" (p. 28). Causal attributions may hold the key in finding the underpinning mechanism that connects teacher racial attitudes and behaviors, especially disciplinary, toward their students of color.

Causal Attributions and Racial Attitudes

Causal attribution focuses on the inferring of the *why* of others' actions (Eagly & Chaiken, 1993). Causal attributions can ultimately lead to negative attitudes toward students (Weiner, 2000), and the race of the student may play a role in these attributions. Researchers have found that teachers often treat children of color differently in the classroom, exhibiting negative attitudes and low expectations (Little & Welsh, 2022; Papageorge et al., 2020). These are communicated both verbally and nonverbally, as well as through increased discipline, which may

in turn have a negative effect on both short- and long-term educational outcomes (Balfanz et al., 2015; Bates & Glick, 2013; Carter & Goodwin, 1994; Frankenberg, 2012). These negative attitudes may be direct functions of interpersonal causal attributions teachers make regarding student behavior.

Research has consistently shown that teachers hold lower standards for Black students both academically and behaviorally, and that teachers regularly express an unfavorable rating of Black students on behavior, personality, and motivation measures, which in turn affects how they treat students in the classroom (Chang & Demyan, 2007; Chang & Sue, 2003). Chang and Sue (2003) sought to determine if teachers' labeling of behaviors varied by race in stereotypical fashion. Describing undercontrolled behavior as a type of disruptive behavior which appears to be not under the power of the actor, they pointed out that "teachers tend to rate [Black] children higher on undercontrolled behaviors and overall behavior problems than [White] students," and teachers generally attribute Black children's misbehavior to undercontrolled behavior (Chang & Sue, 2003, p. 235). The researchers surveyed 193 teachers (83% female, 74.1% White) with vignettes and found a statistically significant effect in teachers' labeling Black students as predominantly acting out in undercontrolled behaviors (Chang & Sue, 2003). The researchers also found a significant main effect when they examined uncontrolled behaviors on locus of causality, stability, and controllability (Chang & Sue, 2003).

It is also interesting to note that when Chang and Sue (2003) asked teachers what they thought was the primary cause of the child's behavior, 40.9% said personality factors, which by Weiner's (1979) definition are internal, unstable, controllable factors when viewed through the lens of attribution theory. Race did not factor into their explicit explanations for their attributions of behavior; how teachers form these perceptions may be dependent on implicit attitudes (Glock & Krolak-Schwerdt, 2014).

Implicit Attitudes

Implicit attitudes are those that exist in unconscious thought; they drive perceptions and decisions without the person holding them being entirely aware of them. These are automatic responses (van den Bergh et al., 2010), in contrast to explicit attitudes, which are arrived at afterthought and decision.

Teachers may assume that their judgments are not biased, but these assumptions reflect explicit attitudes, which are activated after people have time to think through actions and rationalize them (Glock & Krolak-Schwerdt, 2014). There are five stages to social information processing: encoding, storage, retrieval, inferences, and judgment (Glock & Krolak-Schwerdt, 2014); these stages also take place in teachers as they make attributions about their students' behaviors (Glock & Krolak-Schwerdt, 2014). Explicit attitudes are in the judgment stage of social information processing as opposed to implicit attitudes, which are used in the inferences stage, cognitively activating stereotypes at an earlier stage of processing, one that may be viewed as unconscious (Glock & Krolak-Schwerdt, 2014). Implicit attitudes, which are a direct function of beliefs, lead to behavior that is expressed both verbally and nonverbally toward others.

Implicit attitudes are cognitive functions that affect choices, resulting in specific behaviors (Fishbein, 1966; Yang & Montgomery, 2013). Regarding race, these attitudes may influence teacher behavior that sends messages about how they feel toward their students and what they expect out of their students regarding classroom behavior, educational outcomes, and educational attainment. Research suggests that teachers' implicit racial attitudes also influence how they work with their students in the classroom. Teacher attributions of student behavior may affect these implicit attitudes, but this may also be a bidirectional relationship, where implicit attitudes also affect attributions.

There are a small number of studies that look at implicit attitudes and their impacts on education, even though the research on explicit attitudes is plentiful (Glock & Kovacs, 2013). To this extent, there is a shortage of implicit attitude measures among TES that specifically examine race and how it affects student outcomes, specifically in contexts within the United States (Glock & Kovacs, 2013).

Research Question and Hypothesis

The empirical research on the connection between implicit racial attitudes and causal attributions is slim to date (Glock & Kleen, 2021); however, this literature review has shown that teachers' implicit attitudes affect their high-stakes decisions about students in the classroom. The cognitive process through which this happens has been widely speculated upon; we suggest that teachers' implicit racial attitudes influence their perceptions of and responses to students' classroom behavior specifically through causal attributions. This study addresses the following research question: how are TES' implicit racial attitude scores on an implicit bias test related to their interpersonal attributions of student behavioral challenges in the classroom?

The literature has shown that teachers hold attitudes toward their students of color that, when viewed through the lens of attribution theory, contribute to teachers holding Black students more accountable for their perceived negative actions than their White student counterparts (Battey & Leyva, 2016; Bottiani et al., 2017; Gregory et al., 2010; Milner & Tenore, 2010; Okonofua & Eberhardt, 2015; Skiba et al., 2011). This study will build on these findings by asking if TES who score higher on an implicit racial bias measure will attribute Black students' classroom misbehaviors to lack of control more so than White students. Specifically, it is hypothesized that TES who score higher on the racial implicit bias test will attribute internal causality and controllability to explain challenging behaviors in the classroom more frequently for Black students than for White students.

Method

The sample for this study was 233 teacher education students (age range 19-54 years old, $x=25.63$) in a teacher preparation program situated in a large urban area. Participants were recruited through an online research participation system at the teacher preparation program, located in a college in the northeastern United States; participants received one credit for participation, which was required for their coursework in the semester they participated. Most participants self-identified as female ($n=187$) and participants' self-identified races varied, with 112 participants indicating they identified as White, 63 as Hispanic or Latinx, 48 as Asian, 16 as

Black, 4 as Native American, and 15 as other. When asked for their race/ethnicity, participants were asked to check all that apply, so this *N* totaled more than 233.

Measures

Implicit Associations Test (IAT)

In this study, participants completed the racial bias Implicit Associations Test (IAT) using open-source software. The IAT, originally created by Greenwald et al. (1998), examines associations between concepts, namely between a concept and its relation to either a positive or negative association (Greenwald et al., 2009). The goal of the IAT is to measure implicit attitudes by pairing two concepts and measuring the speed of response (Glock, Krolak-Schwerdt, et al., 2013; Greenwald et al., 1998). The idea is that “the more strongly the participant associates two concepts, the faster the participant will respond when this particular pair of concepts is presented” (van den Bergh et al., 2010, p. 503).

In the IAT procedure (Greenwald et al., 1998), two sets of concepts are chosen, the first being the set of implicit attitudes the researchers desire to study (e.g., White and Black races) and the second being a bipolar attitude representation (e.g., good and bad or positive and negative). Participants are then asked to go through a set of trials categorizing a mixture of the sets (e.g., Black and good; White and good; Black and bad; White and bad) using keyboard keys for their answers. Typically, the keys used are the “I” and the “E” keys. While this exercise varies by study, including the use of practice trials and differing category sets, the typical process remains consistent (Greenwald et al., 2009). Central to the measurement of the IAT are response latencies in milliseconds: “an *IAT effect* is defined as the difference in mean latency between these two conditions (noncompatible minus compatible)” (Greenwald et al., 1998, p. 1468). The IAT is computer-administered to properly assess response latencies (Glock et al., 2013; Greenwald et al., 1998; Greenwald et al., 2009; van den Bergh et al., 2010). In this study, validity was established in several respects; convergent validity, divergent validity, and predictive validity were established by the researchers (Greenwald et al., 1998; Greenwald et al., 2009).

Classroom Behavior Vignettes

Two-sentence long vignettes, which described the classroom misbehavior, were paired with a stock photo of either a Black or a White male child approximately 8-12 years of age. The classroom misbehaviors presented were common misbehaviors rooted in child pathology diagnostic categories: ADHD symptoms, internalization, non-compliance/disruptive behavior, and non-compliance/disorganization. There were eight vignettes total, two per diagnostic category; each pair of vignettes presented one with a Black male stock photo and the other with a White male stock photo, with both vignettes focused on similar behaviors for that diagnostic category. Each vignette was followed by two questions asking the participant to rate attributional dimensions of causality (“the student’s behavior is due to the choices he makes”) and controllability (“the student’s behavior is due to circumstances out of his control”) on a Likert scale of 1 (“not at all”) to 6 (“completely”).

Demographic Questionnaire

Participants completed a demographic questionnaire at the end of the two surveys, which asked about their ages, race/ethnicities, genders, and other background information.

Procedure

First, participants completed the IAT using their computer keyboards, which provided their response latency times in milliseconds (ms). This test was given first to avoid a priming affect by either the vignettes or the demographic questions.

The participants then completed all eight vignettes and the attributional dimension questions, alternating student races while presenting the corresponding diagnostic criteria back-to-back (i.e., first vignette was the White student displaying non-compliance/disorganization criteria; second vignette was the Black student displaying non-compliance/disorganization criteria; third vignette was the White student displaying non-compliance/disruptive criteria; fourth was the Black student displaying non-compliance/disruptive criteria; and so on).

After completion of the vignettes and corresponding questions, participants completed the demographic questionnaire.

Results

First, IAT effect sizes for each participant were calculated using the process developed by Greenwald et al. (2003), discussed in detail below. The effect size was represented by Cohen's d , where greater bias against Black protagonists was indicated by a higher negative effect size. The first analyses that were done were Pearson correlations for participant race and IAT effect size, for participant gender and IAT effect size, and for participant age and IAT effect size to establish baseline statistics. Means and standard errors for the specific attributional dimensions (causality and controllability) asked about in the vignettes were calculated, broken down by race, gender, and age. These analyses were done for descriptive purposes.

To answer the question of whether TES who score higher on the racial implicit bias test attribute internal causality and controllability to explain challenging behaviors in the classroom more frequently for Black students than for White students, difference scores in the vignette questions with matched student races were first calculated. For instance, Question 1 in the two vignette sets is identical except for a change in the represented students' races; these are the difference scores that were calculated, subtracting the rating for the White student from the rating for the Black student. Then these scores were used in a multiple regression using Cohen's d as the regressor on the differences between the fictional vignette Black and White students in identical classroom situations presented to the participants.

IAT Data Review and Effect Size Calculations

First, IAT effect sizes were calculated, beginning with removing data points that were above or below thresholds which indicated false information. Greenwald et al. (1998) considered any response time below 300 ms to be too fast to represent a good faith effort, as the participant selection was most likely done prior to observing the stimulus, and over 3,000 ms to be "momentary inattention" (p. 1467). In their development of the instrument, they recoded anything below 300 ms as the lower threshold and anything above 3,000 ms as the upper

threshold (Greenwald et al., 1998). In contrast, Glock, Kneer, & Kovacs (2013) opted to run smaller margins, throwing out any responses that were under 150 ms and above 1,500 ms.

For this data analysis, the researchers opted to use modified Greenwald et al. (1998) thresholds, marking the lower threshold at 250 ms and the upper threshold at 3,000 ms, combined with Glock, Kneer, and Kovacs' (2013) process of throwing out responses that fell outside of these thresholds. This combination was selected because Greenwald et al.'s (1998) justification for their thresholds made the most sense with the data collected in this study as compared to Glock, Kneer, and Kovacs' (2013) more liberal lower threshold and more conservative upper threshold; however, Glock, Kneer, and Kovacs' (2013) decision to throw out data points that surpassed their threshold rather than recoding to meet the lower and upper bounds meant that trials that did not show good faith efforts at responding would not be counted in the analyses.

The full sample in this study presented a range of responses from 250 ms to 3,000 ms, rating anything below the lower threshold to be a false start and above the upper threshold to be a wandering attention span. This resulted in two participants being removed from the analysis entirely due to more than ten of their responses per trial block being below the lower threshold. Other responses that were removed were on an individual case-by-case basis according to the R script, discussed below.

After reviewing this data and removing IAT selections outside of the specified range, IAT effect size was calculated using R and a script based upon the Greenwald et al. (2003) scoring algorithm. The IAT effect sizes ranged from -1.32 to 0.83 (N=228) and had a mean of -0.11 ($sd=0.36$), with a higher negative effect size indicating a greater bias against Black protagonists and a positive effect size indicating a greater bias against White protagonists.

IAT Correlations

To determine whether study participants' IAT scores were associated with demographic characteristics, first Pearson correlations were calculated for participant race and IAT effect size, for participant gender and IAT effect size, and for participant age and IAT effect size to establish baseline statistics. Table 1 presents these correlations, none of which reached statistical significance, indicating that there were no significant relationships between participant characteristics and their IAT scores.

Table 1

Pearson Correlations: Participant Demographics and IAT Scores

Participant Characteristic (N = 228)	R
Age in years	.03
Gender	.04
Race/Ethnicity	
White	-.18
Black	.29
Hispanic or Latinx	.10

Native American	.01
Asian	-.07
Other	.08

Attributions

Attributional Dimension Means and Standard Errors

Means and standard errors for the specific attributional dimensions of causality and controllability were calculated, as seen in Table 2. To assess locus of causality, participants were given the statement, “The student’s behavior is due to choices he makes,” and asked to rate their agreement on a six-point Likert scale; the higher the ranking, the more participants believed that the student had an internal locus of control. To assess controllability, participants were given the statement, “The student’s behavior is due to circumstances outside of his control,” and asked to rate their agreement on a six-point Likert scale; the higher the rating, the more participants believed that student lacked controllability.

Table 2

Means and Standard Errors of Attributional Dimensions by Vignette Race

Pathology criteria	Student race	Causality		Controllability	
		<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Disorganization	White	2.67	0.09	4.27	0.09
	Black	3.02	0.09	3.94	0.10
Disruptive behavior	White	4.11	0.09	2.76	0.08
	Black	3.94	0.09	2.94	0.09
ADHD	White	2.54	0.08	4.32	0.09
	Black	2.95	0.08	3.90	0.08
Internalization	White	2.46	0.08	4.28	0.09
	Black	2.76	0.09	3.94	0.09

IAT and Attributional Dimension

To answer the question of whether TES who score higher on the racial implicit bias test attribute internal causality and controllability to explain challenging behaviors in the classroom more frequently for Black students than for White students, difference scores in the vignette questions with corresponding behaviors were first calculated between student races. For instance, Question 1 in the two vignette sets is identical except for a change in the represented students’ races; these are the difference scores that were calculated, subtracting the rating for the White student from the rating for the Black student. Then these IAT scores were used in a regression using Cohen’s

d as the regressor on the differences between participants' responses to Black and/versus White students in identical classroom situations.

There were no statistically significant relationships found regarding the IAT predicting difference scores between Black and White students on measures of the attributional dimensions of causality and controllability, as seen in Table 3.

Table 3

Regression of IAT on Difference Scores on Attributional Dimensions

Pathology criteria	Causality			Controllability		
	<i>F</i>	<i>p</i>	<i>df</i>	<i>F</i>	<i>p</i>	<i>df</i>
Disorganization	0.01	0.92	227	0.02	0.88	227
Disruptive behavior	1.06	0.31	227	2.89	0.09	227
ADHD	0.44	0.51	227	0.81	0.37	227
Internalization	0.16	0.69	227	3.47	0.06	227

Two dimensions, however, were of note; specifically, under controllability, disruptive behavior and internalization symptoms came close to a statistically significant effect, indicating that the IAT might be better at predicting these dimensions than in other areas.

Attributional Dimension Mean Comparison

To extend these analyses, paired sample *t*-tests were calculated to determine if contrasting-race vignettes with matching pathology criteria differed in ratings of attributional dimensions, as seen in Table 4.

For the attributional dimension of causality, a higher rating indicated TES belief in greater levels of internal causality; therefore, a positive *t*-score indicated that TES endorsed a higher level of internal causality for the White student than his Black counterpart and a negative *t*-score indicated that TES endorsed a higher level of internal causality for the Black student than his White counterpart.

For the attributional dimension of controllability, a higher rating indicated TES belief in lower levels of controllability; therefore, a positive *t*-score indicated that TES endorsed more controllability for the Black student than for his White counterpart and a negative *t*-score indicated that TES endorsed more controllability for the White student than his Black counterpart.

There were statistically significant differences on all the diagnostic categories. For students with disruptive behavior, participants indicated that they felt that the White student had an internal locus of causality and a lack of controllability for his behavior more than his Black counterpart. For the other three pathology criteria (i.e., disorganization, ADHD, and internalization),

participants believed that the Black students had more internal causality and less controllability than their White counterparts. Table 4 displays these scores.

Table 4

Paired Sample t-Tests for Student Racial Differences on Attributional Dimensions

Pathology criteria	Causality		Controllability	
	<i>t</i> (227)	<i>p</i>	<i>t</i> (227)	<i>p</i>
Disorganization	-3.446	0.001	3.794	0.000
Disruptive behavior	2.579	0.011	-2.704	0.007
ADHD	-5.116	0.000	5.325	0.000
Internalization	-4.603	0.000	4.358	0.000

Additional analyses indicated that there was no significant relationship in the aggregate Black and White student scores between the IAT and both the causality and the controllability measures. In other words, when aggregated across diagnostic categories, there was no significant relationship between IAT scores and attributional dimensions for either set of student scores.

Discussion

This study hypothesized that TES who score higher on the racial implicit bias test will attribute internal causality and controllability to explain challenging behaviors in the classroom more frequently for Black students than for White students. The results of this study did not directly support the hypothesis, although there were findings of note that extended from the hypothesis. Regression analyses found that racial bias IAT scores did not predict attributions of causality or controllability by race of the student for any of the challenging classroom behaviors, indicating that implicit racial bias as measured by the IAT in this study did not have bearing on the differences in assigning attribution to student behaviors described in our vignettes. One important reason for this, discussed further in this section, is that the IAT non-significant scores around the IAT predicting bias may involve questions around what the test for measuring implicit bias actually measures.

There was, however, an important finding regarding TES' interpersonal attributions of Black versus White students who displayed the same challenging classroom behaviors. Paired sample *t*-tests revealed that regardless of IAT scores, TES as a whole were more likely to assign internal causality and controllability to Black students than White students. Using Weiner's (1979) paradigm as a guide, this finding indicates that TES would be most likely to assign ability or mood (depending on stability beliefs, which were not measured in this study) as the interpersonal causal attribution of Black students' behavior when displaying disorganization, ADHD, and/or internalization symptoms. This combination of attributions of internal causality and control might lead TES to feel anger toward Black students, blaming them for their behavior and increasing the likelihood TES would opt for punishment and classroom removal.

Stereotyping may also be a culprit in these decisions. Cultural and racial stereotyping leads teachers to define what are “appropriate” behaviors and what are not, and when students are engaging in “inappropriate behaviors” as defined by the teacher’s dominant cultural referent, the teacher may very well attribute an internal locus of causality and controllability to that behavior, deeming students to be at fault for their choices and leading the teacher to anger, which results in punishment of the offending behavior (see Figure 1 in the literature review). These kinds of behaviors include the very ones that see Black students punished at higher rates than their White counterparts, which may also be indicative of pathology criteria: talking back, loitering, and non-compliance (Gregory et al., 2010; Milner & Tenore, 2010; Skiba et al., 2011). Stereotyping may be leading TES to attribute internal causality and controllability to Black students more than White students because they believe Black students do not willingly act “appropriately.”

Ultimately, the finding that TES were more likely to assign internal causality and controllability to Black students than White students illustrates two important points. First, no matter what their bias as measured by the IAT, TES in this study displayed bias toward Black students who show pathologizing symptoms such as disorganization, ADHD, and internalization through the assignment of more internal causality and less controllability in Black students than their White counterparts with identical pathology symptoms. This tracks with current literature on the topic (Battey & Leyva, 2016; Bottiani et al., 2017; Gregory et al., 2010; Milner & Tenore, 2010; Okonofua & Eberhardt, 2015; Skiba et al., 2011) and indicates that systemic racial bias regarding Black male students runs deep in the United States school system regardless of responses to measures of personal implicit racial bias.

The second point, and the first limitation of this study, is that the IAT may not be a strong enough measure to pick up the effect of implicit bias as a predictor of hypothetical behavior. The literature argues that since cognition is the step prior to behavior, and implicit attitudes – a function of implicit beliefs – are part of that cognitive processing, implicit attitudes therefore affect decision-making and behavior (Glock & Krolack-Schwerdt, 2014). Since implicit attitudes can impact behavior, and the racial bias IAT in this study did not show an effect, the issue may be that latent response time may not actually measure bias. The findings in this study indicate that regardless of racial bias IAT score, TES are more likely to assign controllability and internal locus to Black than White students, which matches the literature quite closely. Additional analyses indicated that when data was aggregated across diagnostic categories, there was no significant relationship between IAT scores and attributional dimensions for either set of student scores, for both Black and White students. What then is the meaning of the lack of relationship with IAT scores? Most likely, the interpretation of IAT score might not be what it claims to be, and therefore the question becomes what latent response time represents. Additionally, because the IAT is intended to measure implicit attitudes and the vignettes were asking for explicit hypothetical behavioral actions, the application of the IAT to hypothetical behavior may not be a direct link. Both questions should be explored in future research.

Limitations

As mentioned above, the first limitation of this study is that there is debate around the IAT and what it actually measures (Greenwald et al., 2019). There are other measures that are currently available that claim to measure implicit attitudes, including priming tests, multiple versions of

the IAT, and variations of the IAT. It may be worth exploring these research questions again with a different measure of implicit racial bias to see if there is one that may work better with hypothetical explicit decision making. This leads to a larger question of what, exactly, is the IAT measuring, which is worth exploring further, particularly with specific populations such as teachers.

One other limitation of note is how the demographics were analyzed; participants were able to select their race/ethnicity from a list of options, and they had the option to collect all that applied. When the data was analyzed, participants were grouped according to the individual races they selected. However, it is important to note that the analyses did not distinguish between White-Hispanic and White-non-Hispanic participants; this is something that should be more closely analyzed in the data set moving forward.

Future Directions

The finding that, regardless of IAT score, TES in the study were more likely to assign internal causality and controllability to Black students than to White students indicates that there is a need for real-world, observed data on how these attributional biases play out in the classroom and how these observations connect to measures of implicit racial bias. These results are consistent with the findings that Black students are more likely to be disciplined for similar infractions than their White counterparts. However, these results were based upon a hypothetical situation presented to participants in a survey as opposed to in-the-moment decision making based on teacher-student relationships that is a hallmark of real-world teaching.

To that end, observational measures are needed as well as a variety of implicit racial bias measures, including and beyond the IAT. Teaching is not a robotic activity; each interaction with a student is embedded in a history of teacher-student interactions, previously obtained knowledge of the student such as student records or disability information, and teacher experience, among other factors. Observational measures in a classroom context will be able to capture some of these relationships that impact interpersonal attributions and ultimately disciplinary decisions at the classroom level. The answer to why teachers choose to discipline Black students at higher rates than their White peers may reside outside of a survey that asks participants for hypothetical decision-making.

Finally, this work should be extended to examine the intersection of race and gender, which was specifically avoided in this research to narrow the focus of the analysis. Future research should include vignettes that examine the same constructs in Black and White females; this work can then be compared with the current data set in order to glean insights into how participants view presentation of diagnostic criteria in a combination of racialized and gendered students.

The sample in this study was based in a large urban city, and one factor that could be explored by using a national sample would be to use the same measures to explore how attributional dimensions are viewed by race of student between TES in urban, suburban, and rural settings.

Conclusion

This study sought to find a connection between TES' implicit racial attitudes and attributions of classroom behavior. Ultimately, the hypotheses were not confirmed, but there was evidence that regardless of implicit bias scores, when presented with symptoms of disorganization, disruptive behavior, ADHD, and internalization, TES were more likely to attribute an internal locus of causality and controllability to Black students than to their White counterparts.

Teacher preparation programs have a minimum standard of preparing TES to work with a diverse set of students, but it is worth interrogating if this is actually happening and, if so, *how* it is happening. If the literature has shown that teachers in the field are not prepared to work with a diverse population of students, it is worth asking if teachers are being prepared at the TES level – before they even walk into a classroom to student teach or, in the case of alternative certification, walk into a classroom to teach for the first time with little, if any, real-world classroom preparation – to interrogate their biases and how these impact their decision making in the classroom. Teacher preparation programs can build in use of the interpersonal attributional theory framework to their curricula, specifically beliefs about the locus of causality and beliefs about controllability, as a framework for this interrogation, including how these beliefs illicit feelings of sympathy and anger as seen in Figure 1 of this literature review (Weiner, 2000, 2007). These beliefs may be explicit or implicit, and it is important that those interrogations begin early and often in programs that prepare teachers for the classroom.

TES are not neutral actors in teacher preparation programs; each comes with knowledge they have gained in their 13+ years' experience as students themselves to inform how they see the students they will soon teach. It is worth asking questions about how these internal beliefs – either implicit or explicit – affect what they learn about managing their expectations around students' behavioral and emotional regulation and to what they attribute these behaviors.

References

- Balfanz, R., byrnes, v., & Fox, J. (2014). Sent home and put off track: The antecedents, disproportionalities, and consequences of being suspended in the ninth grade. *Journal of Applied Research on Children: Informing Policy for Children at Risk*, 5(2), 1-19. <https://doi.org/10.58464/2155-5834.1217>
- Bates, L. A., & Glick, J. E. (2013). Does it matter if teacher and schools match the student? Racial and ethnic disparities in problem behaviors. *Social Science Research*, 42, 1180-1190. <https://doi.org/10.1016/j.ssresearch.2013.04.005>
- Battey, D., & Leyva, L. A. (2016). A framework for understanding Whiteness in mathematics education. *Journal of Urban Mathematics Education*, 9(2), 49-80. <https://doi.org/10.21423/jume-v9i2a294>
- Beckman, L. (1970). Effects of students' performance on teachers' and observers' attributions of causality. *Journal of Educational Psychology*, 61(1), 76. <https://doi.org/10.1037/h0028821>
- Bottiani, J. H., Bradshaw, C. P., & Mendelson, T. (2017). A multilevel examination of racial disparities in high school discipline: Black and White adolescents' perceived equity, school belonging, and adjustment problems. *Journal of Educational Psychology*, 109(4), 532-545. <https://doi.org/10.1037/edu0000155>
- Carter, R. T., & Goodwin, A. L. (1994). Racial identity and education. *Review of Research in Education*, 20, 291-336. <https://doi.org/10.3102/0091732X020001291>
- Castro, A. J. (2010). Themes in the research on teacher education students' views of cultural diversity: Implications for researching millennial teacher education students. *Educational Researcher*, 39(3), 198-210. <https://doi.org/10.3102/0013189X10363819>
- Chang, D. F., & Demyan, A. (2007). Teachers' stereotypes of Asian, Black, and White students. *School Psychology Quarterly*, 22(2), 91-114. <https://doi.org/10.1037/1045-3830.22.2.91>
- Chang, D. F., & Sue, S. (2003). The effects of race and problem type on teachers' assessments of student behavior. *Journal of Consulting and Clinical Psychology*, 71(2), 235-242. <https://doi.org/10.1037/0022-006X.71.2.235>
- Cooper, E., & Allen, M. (1998). A meta-analytic examination of the impact of student race on classroom interaction. *Communication Research Reports*, 15(2), 151-161. <https://doi.org/10.1080/08824099809362109>
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.
- Fishbein, M. (1966). The relationships between beliefs, attitudes and behavior. In S. Feldman (Ed.), *Cognitive consistency, motivational antecedents and behavioral consequents* (pp. 199-223), Academic Press.
- Frankenberg, E. (2012). Exploring teachers' racial attitudes in a racially transitioning society. *Education and Urban Society*, 44(4), 448-476. <https://doi.org/10.1177/0013124510392780>
- Gay, G., & Kirkland, K. N. (2003). Developing cultural critical consciousness and self-reflection in pre-service teacher education. *Theory into Practice*, 42(3), 181-187. https://doi.org/10.1207/s15430421tip4203_3
- Glock, S., & Kleen, H. (2021). The role of preservice teachers' implicit attitudes and causal attributions: a deeper look into students' ethnicity. *Current Psychology*, 1-11. <https://doi.org/10.1007/s12144-021-02000-2>

- Glock, S., Kneer, J., & Kovacs, C. (2013). Teacher education students' implicit attitudes toward students with and without immigration background: A pilot study. *Studies in Educational Evaluation, 39*, 204-210. <https://doi.org/10.1016/j.stueduc.2013.09.003>
- Glock, S. & Kovacs, C. (2013). Educational psychology: Using insights from implicit attitude measures. *Educational Psychology Review, 25*, 503-522. <https://doi.org/10.1007/s10648-013-9241-3>
- Glock, S., & Krolak-Schwerdt, S. (2014). Stereotype activation versus application: How teachers process and judge information about students from ethnic minorities and with low socioeconomic background. *Social Psychology of Education, 17*(4), 589-607. <https://doi.org/10.1007/s11218-014-9266-6>
- Glock, S., Krolak-Schwerdt, S., Klapproth, F., & Bohmer, M. (2013). Beyond judgment bias: How students' ethnicity and academic profile consistency influence teachers' tracking judgments. *Social Psychology of Education, 16*, 555-573. <https://doi.org/10.1007/s11218-013-9227-5>
- Graham, S. (1988). Can attribution theory tell us something about motivation in Blacks? *Educational Psychologist, 23*(1), 3-21. https://doi.org/10.1207/s15326985ep2301_1
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Associations Test. *Journal of Personality and Social Psychology, 74*(6), 1464-1480. <https://doi.org/10.1037/0022-3514.74.6.1464>
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*(2), 197. <https://doi.org/10.1037/0022-3514.85.2.197>
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*(1), 17-41. <https://doi.org/10.1037/a0015575>
- 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. <https://doi.org/10.3102/0013189X09357621>
- Hunter, M., & Barker, G. (1987). "If at first...": Attribution theory in the classroom. *Educational Leadership, 45*(2), 50-53.
- Lacoe, J., & Steinberg, M. P. (2019). Do suspensions affect student outcomes? *Educational Evaluation and Policy Analysis, 41*(1), 34-62. <https://doi.org/10.3102/0162373718794897>
- Little, S. J., & Welsh, R. O. (2022). Rac(e) ing to punishment? Applying theory to racial disparities in disciplinary outcomes. *Race Ethnicity and Education, 25*(4), 564-584. <https://doi.org/10.1080/13613324.2019.1599344>
- Milner IV, H. R., & Tenore, F. B. (2010). Classroom management in diverse classrooms. *Urban Education, 45*(5), 560-603. <https://doi.org/10.1177/0042085910377290>
- National Center for Education Statistics. (2023). Characteristics of Public School Teachers. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. Retrieved July 2023, from <https://nces.ed.gov/programs/coe/indicator/clr>.
- Okonofua, J. A., & Eberhardt, J. L. (2015). Two strikes: Race and the disciplining of young students. *Psychological Science, 26*(5), 617-624. <https://doi.org/10.1177/0956797615570365>
- Papageorge, N. W., Gershenson, S., & Kang, K. M. (2020). Teacher expectations matter. *Review of Economics and Statistics, 102*(2), 234-251. https://doi.org/10.1162/rest_a_00838

- Skiba, R. J., Horner, R. H., Chung, C. G., Rausch, M. K., May, S. L., & Tobin, T. (2011). Race is not neutral: A national investigation of African American and Latino disproportionality in school discipline. *School Psychology Review, 40*(1), 85-107. <https://doi.org/10.1080/02796015.2011.12087730>
- 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. <https://doi.org/10.3102/0013189X20912758>
- van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M., & Holland, R. W. (2010). The implicit prejudiced attitudes of teachers: Relations to teacher expectations and the ethnic achievement gap. *American Educational Research Journal, 47*(2), 497-527. <https://doi.org/10.3102/0002831209353594>
- Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology, 71*(1), 3-25. <https://doi.org/10.1037/0022-0663.71.1.3>
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92*(4), 548-573. <https://doi.org/10.1037/0033-295X.92.4.548>
- Weiner, B. (2000). Intrapersonal and interpersonal theories of motivation from an attributional perspective. *Educational Psychology Review, 12*(1), 17-30. <https://doi.org/10.1023/A:1009017532121>
- Weiner, B. (2007). Examining emotional diversity in the classroom: An attribution theorist considers the moral emotions. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in Education* (pp. 75-88), Academic Press. <https://doi.org/10.1016/B978-012372545-5/50006-X>
- Weiner, B. (2012). An attribution theory of motivation. In P. A. M. Van Lange, A. W., Kruglanski, & E. T. Higgins (Eds.), *Handbook of Theories and Social Psychology* (pp. 135-155), Sage Publications Ltd.
- Yang, Y., & Montgomery, D. (2013). Gaps or bridges in multicultural teacher education: A Q study of attitudes toward student diversity. *Teaching and Teacher Education, 30*, 27-37. <https://doi.org/10.1016/j.tate.2012.10.003>