An Examination of Factors Hypothesized to Moderate Stereotype Threat Effects on the Ravens Advanced Progressive Matrices

by

Jimmy Louis Johnson

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
May 14, 2010

Approved by

Daniel Svyantek, Chair, Professor of Psychology
Bryan Edwards, Assistant Professor of Psychology
Jacqueline Mitchelson, Assistant Professor of Psychology

Abstract

The present study examined the moderating effects of racial identity centrality and performance goal orientation on the relationship between stereotype threat and test performance. I also assessed the extent to which test-taking motivation, test-taking anxiety, fairness, and perceived job-relatedness mediated the stereotype threat-test performance relationship. African-Americans were assigned to one of three conditions, all varying in the amount of associated stereotype threat, and given a cognitive ability test. Information concerning their individual racial identities, goal orientations, and perceptions of tests was collected. Results indicated that stereotype threat was not a statistically significant predictor of test performance. Additionally, none of the proposed moderators and mediators was statistically significant.

Table of Contents

List of Tables

List of Figures

Introduction

The "achievement gap" between Black and White test-takers is the discrepancy in scores on cognitively-loaded tests of knowledge, ability, and achievement between Blacks and Whites. It has been an area of public concern for many years because scores on cognitively-loaded tests are predictive of so many important aspects of personal and professional life (e.g., income, schooling, quality of life). The importance of explaining and subsequently reducing this gap is paramount. Previous literature pioneered initially by Steele & Aronson (1995) suggests that stereotype threat accounts for significant differences on measures of cognitive ability. Stereotype threat effects operate when tests are used for evaluative purposes such as high-stakes testing situations (Steele & Aronson, 1995; 1998). Under these circumstances, the threat of confirming a negative stereotype may have an adverse affect on performance in the stereotyped domain. It is therefore imperative that we identify and understand variables that may moderate the effects of stereotype threat.

The primary objective of the present study is to examine the moderating effects of goal orientation and racial identity on the relationship between stereotype threat (Steele & Aronson, 1995) and test performance. To better understand how stereotype threat affects test performance, I will also investigate the hypothesis that the relationship between stereotype threat and test performance is mediated by test-related perceptions (e.g., anxiety, motivation, face validity). Figure 1 (See APPENDIX A) depicts the proposed model tested in the current study.

1

In the education, testing, and personnel selection literature, tests of ability and achievement have been shown to be the most valid predictors of school, job and training performance (Kuncel, Hezlett, & Ones, 2001; Schmidt & Hunter, 1998). However, it has also been extensively documented that tests of intelligence, ability, and achievement generally display large subgroup differences (Bobko, Roth & Potosky, 1999; Hartigan & Widor, 1989) with a widely-cited one standard deviation difference in Black–White performance (Chan & Schmitt, 1997; Roth, Bevier, Bobko, Switzer, & Tyler, 2001; Schmitt, Clause, & Pulakos, 1996).

This one standard deviation Black–White difference on paper-and-pencil tests of ability and achievement typically results in adverse impact against Blacks. A test or assessment tool displays adverse impact if there are differential outcomes associated with the use of the test (e.g., selection, promotion) as a function of a protected class status variable (e.g., race, sex, color, religion, national origin, age, and disability). The presence of subgroup differences and associated adverse impact has important implications for individuals, organizations and universities, and society at large. For individuals adversely impacted by a test, there is a loss of employment and education opportunities and all the benefits associated with such. From the organization or university's perspective, there is the dilemma of using the most valid predictors of job and academic performance and concurrently minimizing the legal, ethical, and professional liabilities associated with using tests that display subgroup differences (Sackett, Schmitt, Ellingson, & Kabin, 2001). At the societal level, there are issues pertaining to addressing past social wrongs, diversity, and equal opportunity (Doverspike, Taylor, & Arthur, 1999). The racial and ethnic composition of society's workforce and hence socioeconomic strata are directly affected by decisions based on high-stakes testing. For example, standardized achievement tests such as the Scholastic Assessment Test (SAT) are used for determining

admissions and scholarship allocations for colleges and universities. Licensing and certification exams and employment tests, which measure knowledge, skills, and abilities also influence a society's workforce (Sackett et al., 2001). Thus, understanding the mechanisms that promote subgroup differences on selection tests has unquestionable value.

Many theories have been offered to explain observed Black-White differences on tests of knowledge, skill, ability, and achievement (e.g., Arthur & Doverspike, 2003; Ceci, 1991; Ceci & Williams, 1997; Hernstein & Murray, 1994; Helms, 1992; Neisser et al., 1996; Reynolds & Brown, 1984; Rushton, 2000; Sackett et al., 2001). Among the most inflammatory of these explanations is that these differences reflect actual group differences in mental ability that are not only innate, but extremely resistant to change (e.g. Gottfredson, 1988). A more notable contribution to this line of thinking can be seen in Hernstein and Murray's 1994 publication of *The Bell Curve: Intelligence and Class Structure in American Life* in which the authors suggest that our society is becoming stratified by mental ability. Accordingly, the observed differences in cognitive ability also give rise to racial differences in social class suggesting that Blacks are overrepresented in the lower socioeconomic status classes because they are of lower average intelligence than Whites. Using this explanation, a simple solution could be to adjust the distribution of Black test-takers upwards by a constant to match the distribution of White test-takers, particularly in cases that involve the use of these scores for selection in university or organizational settings. Not only was this practice of "race-norming" abolished by the Civil Rights Act of 1991, but it also instills and reinforces the same group differences that it is attempting to extinguish (Gottfredson 1994). Other researchers (e.g., Berry, 2003) propose that the disparities in academic performance arise from differential preparation as a function of the types of schools that children from each group attend. Black students may receive instruction

3

(specifically in the area of mathematics) that opposes their cultural learning preferences, as it may not include techniques such as contextual problem solving, reasoning and proof through individualistic thinking, or concrete examples of abstract concepts (Berry, 2003). Often these differences in instruction are also a function of socioeconomic status differences between students at different schools. One recent theory that shows promise in explaining variance in observed subgroup differences in test scores is stereotype threat.

Stereotype Threat

The foundation for stereotype threat comes from social identity theory (Tajfel & Turner, 1979; 1986). This theory was originally developed to understand the psychological basis of intergroup discrimination. According to the theory, we naturally categorize ourselves and others into groups depending on which characteristics are most salient. These groups can involve political affiliation (e.g., Democrat or Republican), religion (e.g., Muslim or Jew), race (e.g., African American or Caucasian), and many others. Further, we tend to associate with those that are part of the same groups to which we belong (our ingroups). The reinforcement of being surrounded by those that share similar ideals and values boosts self-esteem, and also solidifies one's identity within a particular group. Additionally, people are likely to compare themselves to other social groups in terms of whichever group characteristics happen to be most salient. Not surprisingly, comparers tend to find favorable biases for the ingroup, independent of whether or not those biases are well-founded. Finally, in a similar vein to the comparison component, people wish to be both distinct from, and positively compared to other social groups. Consequently, group members are likely to feel threatened by unfavorable comparisons of their group to other outgroups.

Identification with one's ingroup also predicts preferences for identity management strategies that may be used to counteract a "negative social identity", which results from an ingroup member's comparison of his or her group with a superior or more privileged group. When such comparisons occur, group identification depends on the perception of status inferiority as legitimate or stable, and the boundaries between groups as being permeable or impermeable (Mummendey, Kessler, Klink, & Mielke, 1999). Identity management strategies are individual mobility, recategorization at a higher level, social competition, realistic competition, preference for temporal comparison, and reevaluation of the material dimension (Blanz, Mummendey, Mielke, & Klink, 1998; Mummendey et al., 1999). It is likely that individuals attempting to improve ingroup evaluations that are also highly identified with their ingroup might engage in social or realistic competition with the outgroup competitor (in an attempt to reverse the status relation). In contrast, those who do not identify as strongly with their ingroup may manage their positive social identity by recategorizing themselves at a level that allows them inclusion into the outgroup (e.g., Americans versus Blacks or Whites), or by leaving the ingroup in favor of the superior outgroup altogether.

In short, social identity theory focuses on the cognitive self-aspects of belonging to a group. According to the theory, one starts with the assumption of a "negative social identity" for his or her ingroup and attempts to counteract this position by using a number of different identity management strategies. The theory has strong implications for stereotype threat as a construct, because individuals that are typically affected by stereotype threat have a negative social identity relative to the group to which they are being compared (depending largely on the stereotyped domain in question).

Stereotype threat is a construct that focuses mostly on applicant cognitions during the testing procedure (Steele & Aronson, 1995, 1998). It explains the behavior and outcomes of persons in situations in which a widely-known, social stereotype is highly salient. In the case of Black-White test score differences, the widely-held stereotype is that scores on tests of cognitive ability are lower for Blacks than Whites. The threat of confirming the negative stereotype induces fear that negatively affects test performance. Accordingly, there are expected Black-White differences on standardized tests of knowledge, skill, ability, and achievement because the threat of possibly confirming the negative stereotype elicits anxiety in Black test-takers that can have an adverse effect on performance. Furthermore, the test taker may not internalize or accept the veracity of the stereotype, but the knowledge that it exists for his or her specific subgroup is enough to stimulate the threatening condition. For example, Steele and Aronson (1998) found that presenting a "difficult test as diagnostic of ability" produced enough threat in academically successful Blacks at Stanford University to disrupt performance on an ability test. Research shows that under "threat" conditions, Blacks spend more time answering fewer questions (Steele & Aronson, 1995). Of the questions answered, a larger percentage was answered incorrectly than for White test-takers. Their findings support the idea that the Black test-takers experienced instances of inefficient mental processing as a result of focusing on the the self-significance of the inability to immediately identify the correct answer rather than concentrating on answering the item (Steele & Aronson, 1995). Other researchers (e.g., Sackett, Hardison, & Cullen, 2004) caution against prematurely generalizing these laboratory results to real-world high-stakes testing environments. Steele and Aronson (1995) demonstrated that stereotype threat influences test scores when both ability diagnosticity (e.g., when tests are presented as diagnostic of one's ability) and race are primed in testing situations. However, in most cases, it is impossible to

develop a true non-diagnostic testing situation and not prime race differences. For instance, demographic information, including race is collected before tests are administered, thus priming race. In addition, it is assumed that the stereotype of race differences is widely-known, thus is highly salient for high-stakes tests such as the SAT. As such, the challenge is to develop a high-stakes testing situation in which race and test diagnosticity are not primed, thus preventing stereotype activation. This concern has been previously addressed in laboratory settings by simply designating a test as non-diagnostic of mental ability (Steele & Aronson, 1995, 1998; McKay, Doverspike, Bowen-Hilton, & Martin, 2002). Unfortunately, convincing participants that a high-stakes test is not diagnostic of cognitive ability is both unlikely and potentially unethical.

In the past, researchers have attempted to reduce the effects of stereotype threat in applied settings. This is a very tall order, considering that the mere mention of a test may be enough to activate the negative stereotype associated with performance and subsequently elicit stereotype threat (Steele & Aronson, 1995). Steele et al. (2002) describe an early 1990s dormitory-based intervention at the University of Michigan aimed at reducing stereotype threat effects in real-world school situations. The program involved freshmen students who participated in a serious of late-night discussion groups aimed at reducing stereotype threat. The Black students in the program substantially outperformed Blacks in the general population of the University of Michigan and were less likely to drop out of school in subsequent years. Most importantly, results showed that the more discussions they attended, the less likely they would later report having experienced stereotype threat. In turn, the less the students experienced stereotype threat, the higher their subsequent grades. The results show the utility of educating individuals that may be the target of negative stereotypes (Blacks in this case) about the dangers

of stereotype threat. However, in many situations such as employment testing, providing extra tutoring or guidance may not be practical.

Additionally, Stricker (1998; Stricker & Ward 1998) attempted to reduce stereotype threat effects in applied settings, specifically computerized tests and high school examinations, by comparing the performance of students who were required to report their ethnicity and sex before the high-stakes test against the performance of students who reported their ethnicity and sex after completing the test. The reasoning behind this manipulation was that collecting information concerning race or gender could possibly activate the negative stereotype associated with the testing situation (e.g. Blacks do not perform as well as Whites; women do not perform as well as men). The manipulation was performed for two types of tests: an Advanced Placement Calculus examination as well as a computerized placement test. Advanced Placement courses are college-level classes offered to high school students for which they can potentially earn college credit. In both cases, results indicated that this manipulation did not significantly affect the test performance of Blacks females, or other subgroups of examinees to which a negative stereotype concerning cognitive ability might apply. In other words, collecting demographic information after the test was completed did not affect the degree to which stereotype threat was activated and scores were similar regardless of whether demographic information was collected before or after the test. Unfortunately, the results alone do not indicate that the similar scores were due to the fact that no threat was present, or that both groups experienced the same amount of threat regardless of the manipulation. Assuming that threat was in fact still present, the results demonstrate the need for a more effective method for reducing stereotype threat.

Arguably the most practical and widely-applicable manipulation for reducing stereotype threat effects was performed by Brown and Day (2006). To measure performance, the authors

used the Raven's Advanced Progressive Matrices (APM): a test of abstract ability regarded by many as a "culture-free measure of intelligence." The procedure involved three conditions: one designating APM as an IQ test (high threat), one designating the APM as a measure of observation and clear thinking, although the word "test" is mentioned several times during the instructions (standard threat), and one designating the APM as a set of puzzles (low threat). Results showed that Blacks in the group that designated the APM as a set of puzzles reported lower levels of stereotype threat and had higher levels of test performance. This methodology has been chosen for the current study because tests can be ethically described as puzzles or a measure of pattern recognition, which does not carry the burden of the negative stereotype. Also, it is necessary to replicate past research in order to study the effects of my proposed moderators and mediators.

With this in mind, the procedure will include a non-diagnostic condition (the APM is described as a set of puzzles), and two stereotype threat conditions: stereotype threat light (the APM is described as a measure of perceptual organization) and stereotype threat heavy (the APM is described as a test of cognitive ability). These conditions essentially represent a replication of the conditions used by Brown and Day (2006), with the exception that the stereotype threat light condition in the present study does not use the word "test", whereas the standard threat condition for the Brown and Day study does. By excluding the word "test" altogether, the results should indicate if using the designation "measure of perceptual organization" elicits a similar amount of threat to using the word "test." The conditions were chosen because of the applied focus of the Brown and Day study, which recognizes that manipulations such as designating the test as non-diagnostic of ability (e.g., Steele & Aronson, 1995) are usually not possible. The implication is that in real-world high-stakes testing

9

situations, it becomes difficult to avoid priming race and the diagnosticity of the test and thus

avoiding the negative effects of stereotype threat. Therefore, any effects found in the stereotype

threat light condition should be generalizable to realistic testing situations since the condition is a

subtle and plausible manipulation. Although the relationship between test diagnosticity and

performance has been tested in numerous previous studies, it is necessary to test this hypothesis

before proceeding with subsequent hypotheses which serve as the basis for the primary

objectives. Accordingly, I propose the following:

> *Hypothesis 1a:* Test performance will be higher in the control (no-stereotype threat)
> condition than in the stereotype threat heavy condition.
> *Hypothesis 1b:* Test performance will be higher in the control (no-stereotype threat)
> condition than in the stereotype threat light condition.
> *Hypothesis 1c:* Test performance will be higher in the stereotype threat light condition
> than in the stereotype threat heavy condition.

Moderating Effects

Stereotype threat does not explain all variance in Black-White test score differences, so

research has turned to the search for moderators of the race/test score difference relationships.

For example, some research has shown that there are important situational and individual

difference variables that moderate the relationship between stereotype threat and test

performance. That is, several boundary conditions have been proposed which constrain the

emergence and consequences of stereotype threat. Steele, Spencer, and Aronson (2002) provide a

short list of both situational and individual difference moderators. For example, designating a

test as not diagnostic of ability is likely to reduce the harmful effects of stereotype threat (Steele

& Aronson, 1995; Steele et al., 2002), although this solution is neither realistic nor ethical in

high-stakes testing situations. Task difficulty and frustration are two additional related examples

of situational moderators. In the case of Black participants, difficult tests are often necessary to

10

arouse threat because the negative stereotype will either be instantly refuted by superb performance on easy tests, or may not come to mind at all.

Individual difference moderators that have been investigated in prior research include the following: (a) the degree to which an individual considers himself or herself competent within the specified domain (e.g., domain identification); (b) the degree to which an individual classifies himself or herself as a member of the specified group; and (c) the degree to which an individual is aware of the pervasiveness of the stereotype (e.g., stigma consciousness). These examples and others have been shown to significantly affect the relationship between stereotype threat and test performance, as reporting high levels of any of these three variables have resulted in heightened perceptions of stereotype threat (Steele et al., 2002). Although the moderators reviewed above explain variance in test performance, there are two other variables that should play a crucial role in moderating the stereotype threat/test score relationship: racial identity and goal orientation. I predict that racial identity will be a strong moderator in the present study because the stereotype in question is targeting race. Therefore, an individual's susceptibility to stereotype threat in this case is contingent upon the degree to which that person identifies with the racial group that is targeted by the stereotype. Goal orientation is a construct that has not yet been proposed as a potential moderator, but should prove to be useful for identifying which individuals will experience heightened feelings of stereotype threat. The negative stereotype in question predicts that Blacks will perform significantly worse than their White counterparts on tests of cognitive ability. Because this stereotype involves a specific pattern of performance outcome, individuals with a goal orientation that is chiefly concerned with performance will likely experience stereotype threat to a much greater degree than those who are not as driven by a performance goal orientation.

Racial Identity

As mentioned above, one variable that has been found to change the way stereotype threat impacts performance is the extent to which one identifies with the stigmatized group (e.g., Steele, Spencer, & Aronson, 2002; Davis, Aronson, & Salinas, 2006; McFarland, Lev-Arey, & Ziegert, 2003). For Blacks, this construct is referred to as Black racial identity. Racial identity is conceptualized as one's sense of collective identity based on a shared heritage with other group members (Helms, 1990). Social identities manifest themselves as perceived similarities in attitudes, behavior, and social experiences. Some theorists contend that strong racial identity makes Blacks more susceptible to the stigmas associated with their group membership (Fordham & Ogbu, 1986). Consequently, Steele and Aronson (1995; 1998), in their early explications of stereotype threat phenomenon, found that those Blacks who reported high levels of group identity had heightened experiences of stereotype threat. Other researchers (Oyserman, Harrison, & Bybee, 2001), however, have argued for the opposite effect: that identifying with the stereotyped group might buffer a person from stereotype threat effects. Racial identity may serve as a buffer because individuals who are secure with their racial identity may focus more on positive racial stereotypes instead of negative racial stereotypes. This was the case for a small population of Black girls who were strongly identified with being Black, and were less susceptible to stereotype threat because they replaced the popular negative stereotype with a positive stereotype, thus eliminating the potential threat of the negative stereotype (Ambady, Shih, Kim, & Pittinsky, 2001). The results showed that those most strongly identified with being Black in the sense of believing that being Black is associated with achievement showed greater achievement efficacy than Black girls who were not strongly identified with their racial identities. Additionally, differences in personal encounters with discrimination or negative

12

stereotyping may affect perceptions of stereotype threat, since the stereotype must be sufficiently ubiquitous to arouse threat. In other words, the more widespread a negative stereotype is, the more likely targets are to be conscious of it and subsequently threatened by it.

Previous research (Davis et al., 2006) has examined racial identity as a possible moderator, but only in low-threat conditions (e.g., race was not primed, meaning that there was a much lower potential for negative stereotype activation and associated threat). Low threat conditions are typically not found in real-world high-stakes testing situations. This begs the question of whether moderating effects would still be observed under conditions of high threat which are more common in high-stakes testing. Davis et al. (2006) found that students who strongly internalized racial identity attitudes correctly solved more problems than students who did not internalize racial identity attitudes. The problem with these findings is simply that in these conditions, the stereotype was not activated (with the exception of the students that may deem any test whatsoever as threatening to the self). To assess the generalizability of Davis et al.'s findings, it would be necessary to examine racial identity among test takers who are made aware of the negative stereotype (through the activation of stereotype threat). Therefore, I have chosen to study racial identity in the present study because previous studies did not appear to go far enough in providing evidence that racial identity affects the relationship between stereotype threat and performance. Yet, among all the moderators, racial identity is theoretically the strongest moderator of the stereotype threat-test performance relationship.

Some studies (e.g., McFarland et al., 2003) have incorporated Helms' (1990) uni-dimensional theoretical approach to simply categorize individuals as high or low along the dimension of racial identity. However, current theorizing views racial identity as a very intricate and multi-faceted construct (Sellers, Rowley, Chavous, Shelton, & Smith, 1997). As such, the

13

present research study contributes to the literature by incorporating a multi-dimensional

approach of racial identity in order to more adequately capture the complexity of racial identity

for Blacks and its effect on stereotype threat. The Multidimensional model of racial identity

(MMRI) (Stryker & Serpe, 1994; 1998) defines different aspects of racial identity using the

concepts of identity theory, arguing that the choices that an individual makes are, in part, a

function of the extent to which they are related to a personally relevant role-identity. In essence,

we all have a number of social identities that are ordered hierarchically with respect to how

integral each identity is to the individual's own self-perception and in a given situation (Stryker

& Serpe, 1994). Although the salience of each racial identity may change with each different

situation, the stereotype in the present study involves one very specific situation: high-stakes

tests of cognitive ability. In the case of the present study, the stereotype with which we are

concerned is a racial stereotype. Therefore, in general I posit that individuals with a weak Black

identity are less likely to believe that they are being targeted by the stereotype, thus reducing the

experience of stereotype threat. The MMRI consists of four dimensions: centrality, ideology,

regard, and salience (Sellers et al., 1997).

Centrality refers to the degree to which individuals define themselves with respect to

their race. That is, this dimension provides a measure of how *central* an individual's race is to his

or her self-constructed identity. Accordingly, centrality could represent the difference between

being a *Black* lawyer, or a lawyer who *happens to be Black.* Ideology is the individual's personal

beliefs about the appropriate social conduct for members of his or her race. There are four

proposed ideologies in the MMRI: (a) nationalist, identified by the belief that Blacks are

important and illustriously distinguished; (b) oppressed minority, characterized by the belief that

Blacks are traditionally and institutionally subjugated, and share many commonalities with other

14

downtrodden groups; (c) assimilationist, identified by the belief that Blacks should strive to

embrace and increase similarities between themselves and other Americans; and (d) humanist,

characterized by the belief in the unity of *all* humans. Regard refers to affective assessment of

his or her race. That is, it represents the extent to which a person views being Black as a reason

for celebration or despair. Additionally, this dimension is divided into two components: private

and public regard (Crocker & Luhtanen, 1990; Luhtanen & Crocker, 1992). Private regard is the

extent to which an individual personally feels positively or negatively about his or her

membership to a racial group, whereas public regard is the degree to which an individual

believes that greater society views his or her group positively or negatively. Research supports

the proposition that outgroups' perceptions influence individuals' views about their ingroup

(Luhtanen & Crocker, 1992). However, Crocker, Luhtanen, Blaine, and Broadnaz (1994) found

that Black students' perceptions about how others evaluated their race were not related to their

own evaluations of their race nor their self-esteem scores, indicating that public regard does not

necessarily predict private regard. That is to say that if high levels of private regard shield an

individual from the effects of stereotype threat (much like the aforementioned example of the

young Black girls that identified with being Black on the basis of a high level of achievement

efficacy), those feelings of private regard should not be subject to manipulation from outside

influences or contrary opinions. The final dimension of the MMRI is salience, which refers to the

degree to which an individual's race is a prominent component of his or her self-concept at a

particular moment in time. These dimensions are viewed as relatively stable across time and

situations and subsequently resistant to change (Sellers et al., 1997).

  Of all the MMRI dimensions, I posit that those most likely to moderate stereotype threat

effects are centrality and private regard. Those individuals who report low levels of racial

centrality will likely be less susceptible to the effects of stereotype threat, because their race

(which is also the race targeted by the negative stereotype) is not an integral component of their

identities. This is a logical conclusion considering the aforementioned proposition that ingroup

identification is a moderator of the relationship between stereotype threat and performance

(Steele et al., 2002). Additionally, individuals with high levels of private regard will most likely

attempt to disprove any negative stereotype since it would oppose their positive beliefs

concerning their race. The inconsistency of the negative stereotype with their high positive

regard for the stereotyped race would likely buffer them from the negative effects of stereotype

threat. Although social identity theory (Tajfel & Turner, 1979; 1986) would predict that those

individuals that report high levels of racial centrality are also likely to report high levels of

private regard (Sellers et al., 1997), the possibility remains that race may be an integral part of an

individual's identity for reasons that are not very rewarding. Therefore, I propose an interaction

between centrality and private regard such that high centrality will lead to subpar performance,

specifically when private regard is low. Accordingly, I propose the following:

> *Hypothesis 2a:* High scores on the centrality dimension of racial identity will be
> associated with lower test performance in stereotype threat conditions than in control
> conditions.
> *Hypothesis 2b:* Individuals with high scores on the private regard dimension of racial
> identity will achieve higher performance than individuals with low levels of private
> regard in the stereotype threat conditions.
> *Hypothesis 2c:* There will be a significant interaction between centrality and private
> regard, such that high centrality will lead to lower test performance than low centrality,
> specifically when private regard is also low.

Goal Orientation

Steele et al. (2002) suggested that the way in which individuals approach the test

moderates the effects of stereotype threat. The construct goal orientation is an approach that has

been investigated with other performance tasks (e.g., Dweck, 1986; Elliot & Harackiewicz, 1996; Elliott & McGregor, 2001; Bell & Kozlowski, 2002). Specifically, people hold either a learning (mastery) or performance orientation towards a task (Dweck, 1986). The learning orientation is distinguished by a desire to increase one's skill and knowledge level, whereas the performance orientation is characterized by the desire to demonstrate skill and knowledge for evaluation from a test, a peer, or any other external source. As expected, both orientations can lead to similar outcomes. The primary difference, however, lies in the underlying mechanisms which achieve the similar outcomes. Individuals either approach a task with eagerness to acquire a new skill (learning orientation), or to demonstrate their expertise to external constituents (performance orientation). Although the two orientations seem to be opposites, they are considered separate dimensions. In fact, people can be high or low on both dimensions simultaneously (Button et al., 1996). It is expected that performance goal orientation will moderate the effects of stereotype threat in this case because the stereotype in question makes a prediction about test performance. Additionally, for test-takers with a performance orientation, ability is the factor that differentiates exceptional performers from poor performers. Because they are unaware of other controllable influences that affect performance, they are more likely to quit an exceedingly difficult test and concede that they are not smart. This stable internal attribution can be devastating in high-stakes testing situations. Thus, an orientation that is not concerned with performance would therefore be more adaptive in testing situations that may elicit stereotype threat.

In general, people are predisposed to subscribe to a particular orientation (Button et al., 1996). However, certain situational elements may influence the type of orientation an individual chooses (Bell & Kozlowski, 2002). In the absence of evaluation or assessment, an individual

may be more likely to adopt a learning orientation instead of a performance orientation. Goal orientation also predicts how an individual responds to task difficulty and potential failure (Bell & Kozlowski, 2002). Those individuals with a learning orientation display what researchers have identified as an *adaptive response pattern*, in part because its associated behaviors promote persistence and eventual success. Individuals with a learning goal orientation are less concerned with performance per se, but instead are focused on mastering the task. As expected, people that demonstrate this response pattern are not as easily frustrated by difficulty or obscurity when approaching a new task. In fact, it would be reasonable to assume that for these individuals, each failure can be seen as useful to the extent that one *learns* from the experience and receives valuable feedback information. As a result, the response pattern of learning orientation typically leads to the pursuit of challenging and difficult material to increase one's knowledge or skill set (Dweck, 1986; Button et al. 1996).

Individuals with a performance orientation generally display a *maladaptive response pattern*, particularly because its behaviors typically lead to withdrawal or avoidance if the individual is unsure of his or her ability for success. Additionally, in the event of a poor performance outcome, the individual is likely to attribute it to a lack of ability. In general, difficult or frustrating tasks are usually avoided since optimum performance is not always a guarantee, particularly when the task is novel. Actual or potential failures present a convenient reason for performance-oriented individuals to completely withdraw from the task at hand. However, the presence of external evaluation may allow for differential appraisals of the same task. For example, a performance-oriented female may view playing the guitar as very complicated and difficult, but she may choose to perform as long as she can be sure that her

peers (who find the guitar to be even more complicated, and are subsequently not as proficient at playing it) rate her performance as exceptional (Dweck, 1986; Button et al., 1996).

Recently, several researchers have attempted to expand upon the dichotomous goal orientation designations. Some argue that a four-factor model may be more appropriate for the construct (Elliot & McGregor, 2001). In the four-factor model, the performance and learning goal orientations are separated into constructs of approach and avoid. Performance approach orientation targets an individual's desire to demonstrate his or her ability to others (e.g., I am motivated by the thought of outperforming my peers). In contrast, performance avoid orientation targets an individual's desire to circumvent poor performance (e.g., my fear of performing poorly is often what motivates me). This is not to say that individuals with a performance avoid orientation avoid tasks altogether, but rather that their performance is geared towards attaining a performance that indicates competency while simultaneously *avoiding* incompetency. It seems that the performance approach orientation can be thought of as a self-esteem booster, whereas the performance avoid orientation can be thought of as a self-esteem defense. Each performance orientation dimension is also focused on external evaluation, which can be very debilitating during testing procedures. In a similar vein, learning goal orientation is also separated into approach and avoid constructs. Learning approach orientation represents the typical conceptualization of learning orientation, in which an individual is less concerned with end-result performance and more concerned with acquiring the necessary skills to improve performance. However, in learning avoid orientation, competence is defined in terms of the absolute requirements of the task or of one's own pattern of attainment, and incompetence is the focal point of regulatory attention. Factor analytic results support the independence of all four

orientations, although the two learning goal orientations are positively correlated, and the two performance goal orientations are positively correlated (Elliot & McGregor, 2001).

As previously stated, learning orientation appears to shield individuals from the negative effects of failure. The person who has less of a fear failure may not be subject to the situational anxiety associated with imminent failure. However, a lack of anxiety does not necessarily reduce the occurrence of failure, so I do not predict that individuals that report high learning orientations will attain higher performance (as opposed to lower levels of anxiety) outcomes than those who report low learning orientations, although learning orientation will still be measured to test this assumption. This can be a useful asset in testing situations as the learning-oriented person views the test itself as an opportunity for personal growth, and not a potential threat to his or her ability. Unfortunately for the performance-oriented individual, every testing situation is merely an opportunity to prove competence or avoid incompetence. Not only is this person less likely to prepare for the test, but he/she would most likely sees intelligence as less malleable. The anxiety elicited by the test most likely depends on the individual's self-perception of ability in that specific domain. In the case of stereotype threat, a performance orientation (either performance approach or performance avoid) may lead an individual to withdraw from the task both cognitively and motivationally as a defense mechanism to avoid appearing incompetent and thus confirming the stereotype. This is not to be confused with withdrawing from the task itself, since in high-stakes testing situations, avoiding the test itself is not an option. Based upon this information, I propose that:

> *Hypothesis 3a*: Performance approach goal orientation will moderate the relationship between stereotype threat and test performance. Specifically, a high performance approach orientation will correspond to lower test performance in the stereotype threat condition than a low performance approach orientation.

Hypothesis 3b: Performance avoid goal orientation will moderate the relationship between stereotype threat and test performance. Specifically, a high performance avoid orientation will correspond to lower test performance in the stereotype threat condition than a low performance avoid orientation.

Mediators

The influence of examinee perceptions on test performance has been increasingly studied in the last 15 years (Ryan & Ployhart, 2000) and driven by findings that various perceptions (e.g., face validity) are related to real outcomes such as test performance or decision to withdraw from the job application process. Some researchers have posited that racial differences in standardized test performance can be attributed to subgroup differences in perceived face validity, fairness, predictive validity, belief in the utility of tests, and self-efficacy (Chan & Schmitt, 1997; Chan, Schmitt, DeShon, Clause, & Delbridge, 1997; Edwards & Arthur, 2007; Gilliland, 1994; Ryan, 2001). Ryan (2001) provided the most comprehensive review of the testing literature that attempts to explain race-based test score differences on cognitively-loaded tests through processes of test perceptions. She presents a model in which test perceptions mediate the relationship between race and test performance. She argues that negative test perceptions introduce test-irrelevant cognitions, negative affect, and lowered motivation that interfere with test-relevant behaviors or cognitions (Ryan, 2001). To the extent that tests elicit different perceptions by race, then these subgroup differences in test perceptions could be manifested in subgroup differences in test performance (Edwards & Arthur, 2007).

It may be that stereotype threat decreases test performance for Blacks because the threatening condition elicits negative perceptions related to the test and/or testing situation. Although theory points to test perceptions as an explanatory mechanism for stereotype threat effects, I found only three studies that have examined the relationship between stereotype threat

and test perceptions (Nguyen, O'Neal, & Ryan, 2003; Mayer & Hanges, 2003; Ployhart, Ziegert, & McFarland, 2003). In one study, stereotype threat and test perceptions were moderately and uniquely related to test performance but the mediating role of test perceptions was not assessed (Mayer & Hanges, 2003). Ployhart et al. (2003) obtained evidence that test-taking perceptions (e.g., anxiety, face validity, and motivation) were significant mediators of the race-test performance relationship and that this relationship was stronger in stereotype threat conditions.

These three studies have been criticized on a number of points (Steele & Davies, 2003; Steele & Aronson, 2004). The primary criticism is that the studies did not use a true control condition since the word "test" was used in the instruction set to describe the performance task. Steele asserted that the use of the word "test" is enough to elicit stereotype threat for Blacks in a high-stakes testing situation so the control conditions in these three studies were in essence, threatening conditions and *not* control conditions. Consequently, the present study addresses the lack of the no-stereotype threat control group that plagued the three previous studies that have examined the relationships among race, stereotype threat, test perceptions, and test performance. Based on previous literature I propose the following hypotheses:

> *Hypothesis 4*: The relationship between perceived stereotype threat and test performance will be partially mediated by (a) test-taking anxiety; (b) test-taking motivation; (c) perceived job-relatedness; and (d) perceived fairness.

Method

Participants

The sample consisted of 183 Black participants at an historically Black southeastern university. Respondents ranged in age from 18 to 54 years with an average of approximately 23 years. Of the 183 participants, 66 were male, 115 were female, and 2 did not report gender. There were 17 freshmen, 36 sophomores, 58 juniors, 67 seniors, 1 graduate student, and 4 participants who did not report classification.

Procedure

Participants were randomly assigned to one of three conditions defined by the strength of stereotype threat as previously used by Brown and Day (2006). In the *stereotype threat heavy* condition, participants were told that the Raven's APM was a test of intelligence. This condition was designed to make salient that the test was diagnostic of ability, which is consistent with other experimental manipulations of stereotype threat reported in the literature (Steele & Aronson, 1995). In the *stereotype threat light* condition, participants were told that the Raven's APM was a test of perceptual organization. The use of the construct label "perceptual organization" to describe the nature of the test in this condition was done so that it would not be explicit to participants that the test was diagnostic of ability. However, the mention that the measure was a test could have been enough to elicit a threatening situation. In contrast to the stereotype threat heavy condition in which mental ability diagnosticity is made salient, the stereotype threat light condition may reduce the threat. Further, the stereotype threat light

condition is an example of a possible, simple manipulation that could be used in real world, high-stakes testing situations to reduce stereotype threat for Blacks and subsequent subgroup differences in test performance. Finally, in the control condition the word "test" or "ability" was never mentioned. Instead, participants were asked to solve a set of puzzles. This condition is consistent with other non-threatening conditions used in the stereotype threat literature (Steele & Aronson, 1995).

All participants were seated in a lecture hall and randomly assigned to one of the three conditions. Next, participants were administered the specific test instructions which corresponded with their assigned conditions and given 15 minutes to complete the Raven's APM. Steele warns that the mere mention of a test or collecting race identification data before taking a test is enough to elicit a threatening condition, thereby nullifying the non-diagnostic instructions in the control condition (Steele & Davies, 2003). Therefore, all test perception (e.g., motivation, face validity) and race-related (e.g., Black racial identity) data were collected after the stereotype threat manipulation condition and Raven's APM were presented. Participants completed the remainder of the measures in the following order: perceived stereotype threat, perceived level of effort in the study, randomized test perception measures (e.g. perceived job relatedness, fairness, test-taking motivation, and test-taking anxiety) goal orientation, and the Multidimensional Inventory of Black Identity.

Measures

Perceived stereotype threat. Several subjective indicators (11 total items) of stereotype threat were included and served as a manipulation check for the experimental conditions: The Perceived Stereotype Threat Scale (Ziegert, Ployhart, & McFarland, 2002), the Post-test Attitudes Survey (McKay, 1999), and the stereotype threat perception measure developed by

Sawyer and Hollis-Sawyer (2003). All ratings were made on a 5–point Likert-type scale (1 = *strongly disagree*; 5 = *strongly agree*) and scores were averaged to create composite scores. The internal consistency estimate (i.e., coefficient alpha) for scores on the perceived stereotype threat scale was .63. Items can be found in Appendix B.

Fairness. Perceived fairness was measured using three items adopted from Smither et al. (1993). Examples of items are "The test results will accurately reflect how well I perform on this test" and "I deserve the test results that I will receive on this test." All ratings were made on a 5–point Likert-type scale (1 = *strongly disagree*; 5 = *strongly agree*) and scores were summed to create composite scores. The internal consistency estimate for scores on the fairness scale was .71.Items can be found in Appendix B.

Perceived Job-Relatedness. Perceived job-relatedness was measured using five items adopted from Smither et al. (1993) such that they applied to college entrance. Examples of items are "Failing to pass this test clearly indicates that you can't pass many college courses," "It would be obvious to anyone that this test is related to college performance," and "I am confident that the test can predict how well an applicant will perform in college courses." All ratings were made on a 5–point Likert-type scale (1 = *strongly disagree*; 5 = *strongly agree*) and scores were summed to create composite scores. The internal consistency estimate for scores on the perceived job-relatedness scale was .80. Items can be found in Appendix B.

Test–Taking Motivation. Test–taking motivation was measured using eight items adopted from the Test Attitude Survey (Arvey et al., 1990; Items from Scale 1). Examples include "Doing well on this test is important to me" and "I tried my best on this test". All ratings were made on a 5–point Likert-type scale (1 = *strongly disagree*; 5 = *strongly agree*) and scores

were summed to create composite scores. The internal consistency estimate for scores on the test-taking motivation scale was .90. Items can be found in Appendix B.

Test–Taking Anxiety. Test–taking anxiety was measured using 10 items adapted from the TAS (Arvey et al., 1990; Items from Scale 4). Examples of items are "I am not good at taking tests" and "I usually get very anxious about taking tests". All ratings were made on a 5–point Likert-type scale (1 = *strongly disagree*; 5 = *strongly agree*) and scores were summed to create composite scores. The internal consistency estimate for scores on the test-taking anxiety scale was .87. Items can be found in Appendix B.

Black Racial Identity. Black racial identity was measured with the Multidimensional Inventory of Black Identity (MIBI; Sellers et al., 1997). The MIBI includes several dimensions and attempts to accommodate the diversity of the Black experience. Of these, the centrality dimension of the MIBI reflects the extent to which one normatively defines him or herself in racial terms. Additionally, private regard can be likened to feelings of pride based on group membership. Centrality was measured using three items from the MIBI, and private regard were measured using four items. All ratings were made on a 5–point Likert-type scale (1 = *strongly disagree*; 5 = *strongly agree*) and scores were created by taking the mean of items on each dimension. The internal consistency estimate for scores on the centrality and private regard scales were .66 and .78 respectively. Items can be found in Appendix B.

Goal Orientation. Goal orientation was measured using the 2 x 2 framework (Elliot & McGregor, 2001). This framework separates the original mastery and performance orientations into mastery approach, mastery avoid and performance approach, performance avoid. A few minor modifications were made to the items to reflect the specifics of this study. The two scales used for this study were performance approach, and performance avoid. Ratings for 6 items (3

for each orientation) were made on a 5–point Likert-type scale (1 = *strongly disagree*; 5 = *strongly agree*). The internal consistency estimate for scores on the performance approach and performance avoid scales were .85 and .80 respectively. Items can be found in Appendix B.

Raven's Advanced Progressive Matrices. *(APM*; Raven, Raven, & Court, 1998). I used the Raven's Advanced Progressive Matrices (APM) Short Form (APM; Arthur & Day, 1994; Raven, Raven, & Court, 1994) which consists of 12 matrix or design problems arranged in an ascending order of difficulty and scored by summing the number of problems answered correctly. The APM short form demonstrates psychometric properties similar to that of the long form with a reduced administration time of 15 minutes. The odd/even split–half reliability with a Spearman–Brown correction for the APM scores was .64.

Results

An initial confirmatory factor analysis was conducted to assess the fit of the measurement

model. Model fit was assessed using the chi–square statistic, the root mean square error of

approximation (RMSEA; Steiger, 1990), the Tucker-Lewis non-normed index (TLI; Tucker &

Lewis, 1973), and the comparative fit index (CFI; Bentler, 1990). The RMSEA is a parsimony–

adjusted index that accounts for model complexity and was used to assess lack of model fit.

RMSEA values less than .05 indicate close approximate fit, values between .05 and .08 indicate

a reasonable error of approximation, and values greater than .10 suggest a poor fit (Browne &

Cudeck, 1993; Hu & Bentler, 1999). The TLI compares the proposed model to the independence

model (e.g. all correlations and covariances are zero) and includes a penalty for adding

parameters. Satisfactory models yield TLI values greater than .90 (Hu & Bentler, 1999). The CFI

assesses the relative improvement in fit compared to the independence model and is resistant to

errors associated with sample size. Satisfactory models yield CFI values greater than .90 (Hu &

Bentler, 1999). The fit of this preliminary model was not acceptable: $\chi^2_{(944)} = 2384.53$, $p < .01$;

RMSEA = .07, (90% CI = .07 – .08); TLI = .66; and CFI = .82. Consequently, a second

confirmatory factor analysis was conducted after removing several problematic items from the

original measures. The fit of the second model was modest: $\chi^2_{(944)} = 1737.98$, $p < .01$; RMSEA =

.07, (90% CI = .07 – .08); TLI = .80; and CFI = .82. This may be due in part to the fact that a

number of items load very highly on multiple factors. However, the items were used to measure

the original dimensions for which they were intended. Therefore, no subsequent changes were made to the model because there was not sufficient rationale to do so.

Descriptive statistics and bivariate relationships are presented for all variables by condition in Table 1, 2, and 3. Data pertaining to levels of perceived stereotype threat was collected as a manipulation check. It was assumed that perceived stereotype threat would be highest in the stereotype threat heavy condition and lowest in the control condition. However, a one-way ANOVA revealed that the differences in perceived stereotype threat across conditions were not statistically significant, $F(2,180) = 1.84$, $p = .162$. Therefore, the manipulation did not appear to work as planned. Hypothesis 1 stated that (a) performance would be higher in the control (no-stereotype threat) condition than in the stereotype threat heavy condition, (b) performance would be higher in the control (no-stereotype threat) condition than in the stereotype threat light condition, and (c) performance would be higher in the stereotype threat light condition than in the stereotype threat heavy condition. Conditions were coded as follows: 1 for control, 2 for stereotype threat light, and 3 for stereotype threat heavy. In addition to providing distinction among groups, the coding also implies magnitude (e.g. stereotype threat heavy = 3 is greater than stereotype threat light = 2). A one-way ANOVA revealed that the differences in performance on the Raven's matrices across stereotype threat conditions were not statistically significant, $F(2, 180) = 2.90$, $p = .06$. In fact, the differences that were observed were in the direction opposite to what was expected. Descriptive statistics indicated that the condition with the highest performance was stereotype threat heavy ($M = 4.97$, $SD = 2.09$) followed by stereotype threat light ($M = 4.56$, $SD = 2.22$) and control ($M = 4.05$, $SD = 1.94$). Tukey's HSD revealed that the only significant difference in performance between groups was between the control and stereotype threat heavy conditions, $t(124) = 2.41$, $p = .05$. The stereotype threat light

condition was not statistically different from either the control condition ($t[110] = 1.28$, $p = .41$) or the stereotype threat heavy condition ($t[126] = 1.07$, $p = .54$). Thus, Hypotheses 1a – 1c were not supported.

Because Hypotheses 1a – 1c were not supported, tests for Hypotheses 2a – 4d were not appropriate, because they are all based on the presupposition of the relationship proposed by Hypothes 1a – 1c. However, because this study was done as a Master's Thesis, the remaining hypotheses were tested anyways for completeness sake. Hypothesis 2 stated that (a) high scores on the centrality dimension of racial identity would be associated with lower test performance in stereotype threat conditions than in the control condition, (b) individuals with high scores on the private regard dimension of racial identity would achieve higher performance than individuals with low levels of private regard in the stereotype threat conditions, and (c) there would be a significant interaction between centrality and private regard, such that high centrality will lead to lower test performance than low centrality, specifically when private regard is also low. Hypotheses 2a and 2b were tested using multiple regression. Hypothesis 2a was tested by regressing test performance onto the stereotype threat condition, racial identity: centrality, and an interaction between the two. Results showed that the interaction term was not statistically significant $F(2, 174) = 1.37$, $p = .26$. Hypothesis 2b was tested by regressing test performance onto stereotype threat condition, racial identity: private regard, and an interaction between the two. Again, the interaction term was not statistically significant, $F(2, 175) = 2.56$, $p = .08$. Hypothesis 2c was tested by regressing test performance onto racial identity: centrality, racial identity: private regard, and an interaction between the two. The interaction term was not statistically significant, $F(1, 176) = .22$, $p = .64$. In sum, because none of the interaction terms of

the three regressions was significant, moderation was not present and Hypotheses 2a – 2c were not supported.

Hypothesis 3 stated that (a) performance approach goal orientation would moderate the relationship between stereotype threat and test performance and (b) performance avoid goal orientation would moderate the relationship between stereotype threat and test performance. Hypothesis 3a was tested by regressing test performance onto stereotype threat condition, goal orientation: performance approach, and an interaction between the two. Results indicated that the interaction term was not statistically significant, $F(2, 175) = .89$, $p = .41$. Hypothesis 3b was tested by regressing test performance onto stereotype threat condition, goal orientation: performance avoid, and an interaction between the two. Again, the interaction term was not statistically significant, $F(2, 175) = 1.86$, $p = .16$. Because none of the interaction terms was significant, moderation was not present and Hypotheses 3a and 3b were not supported.

Hypothesis 4 stated that the relationship between stereotype threat and performance would be partially mediated by four variables: perceived job-relatedness, perceived fairness, anxiety, and motivation. The mediation procedure of Baron and Kenny (1986) and Judd and Kenny (1981) was used to test for mediation. In order to demonstrate *complete* mediation one must first show that the independent variable is predictive of the outcome by running a regression with the independent variable as the predictor and the dependent variable as the outcome. Next, it must be shown that the independent variable is predictive of the mediator by running a regression with the independent variable as the predictor and the mediator variable as the outcome. Third, the mediator variable must be shown to affect the dependent variable by running a regression with both the independent variable and mediator variable as predictors and the dependent variable as the outcome. Finally, to demonstrate full mediation, it must be shown

31

that the relationship between the independent variable and the dependent variable is zero when the mediator is included as a predictor. In order to demonstrate *partial* mediation, it must be shown that the relationship between the independent and dependent variables is significantly smaller than when the mediator is present than when it is not. Accordingly, these procedures were implemented for each of four proposed mediators: anxiety, motivation, perceived fairness, and perceived job-relatedness. As previously addressed in Hypothesis 1, the relationship between stereotype threat and performance was not statistically significant. Additionally, in every case, there was no statistically significant relationship between the predictor (stereotype threat condition) and each mediator: job-relatedness: $F(2, 178) = 1.56$, $p = .21$; perceived fairness: $F(2, 177) = .96$, $p = .39$; motivation: $F(2, 178) = 1.26$, $p = .28$; anxiety: $F(2, 179) = .55$, $p = .59$. Path coefficient values can be found in Table 4. Finally, the Sobel (1982) test statistic was calculated for each of the potential mediator variables. In each case, the test statistic value was not statistically significant. Sobel values and betas can be found in Table 4. In short, the data do not support the presence of mediation for any of the proposed variables.

Discussion

The primary objective of the present study was to examine the moderating effects of goal orientation and racial identity on the relationship between stereotype threat and test performance. With respect to racial identity, it was hypothesized that high levels of centrality would exacerbate the negative effects of stereotype threat whereas high levels of private regard would reduce the negative effects of stereotype threat. Further, high levels of performance goal orientation were also predicted to reduce the negative effects of stereotype threat. Finally, a number of test-related perceptions were proposed to mediate the relationship between stereotype threat and test performance. In short, the results did not support any of the aforementioned propositions.

Although the findings were not statistically significant, the most interesting of the results were the mean differences in test performance across the three stereotype threat conditions. It was expected that performance would be lowest in the stereotype threat heavy condition, and highest in the control (no stereotype threat) condition. Instead, performance was highest in the stereotype threat heavy condition and lowest in the control condition.

There are a number of explanations for this occurrence. First, it is possible that the negative stereotype in question is no longer as widely-held, and is subsequently no longer threatening enough to affect performance. As our society has become more integrated and less segregated, individuals have more opportunities to come into contact with racial outgroup members. These encounters provide firsthand knowledge that may render stereotypes ineffective

and unnecessary, because stereotypes are used primarily in situations in which prior knowledge

is unavailable. Additionally, the possibility exists that the student respondents were inadvertently

coached to resist the negative effects of stereotype threat prior to this study. Historically Black

Colleges and Universities (HBCUs), such as the university at which these data were collected,

typically favor a more collectivist approach to education. Students are often mentored by their

advisors and share a more intimate relationship with their professors than other comparable

institutions. These relationships give professors the opportunity to be candid with their students

about the tribulations of pursuing an academic career as a racial minority, and to prepare them

for instances in which they should expect to be stereotyped or otherwise treated unfairly. There is

also the possibility that the stereotype threat manipulation was not salient enough to either

activate threat, or to distinguish among the three conditions. Although Steele and Aronson

(1995) insist that the mere mention of a measure as a test of cognitive ability can induce threat

and influence performance, this may not be a sufficient trigger at an HBCU. However, the

instructions were written on the test document and not read aloud to the participants, so there

remains the possibility that students may have ignored the instructions that were specific to each

stereotype threat condition and subsequently not been affected by the manipulation. A clustered

random design would have allowed for the specific instructions to have been read aloud to

ensure that the manipulation was recognized. In many cases, Black students that experience

stereotype threat remark that they feel as if their single score is representative of the entire race.

At an HBCU where most all of the students are Black, students are more aware theirs is not the

only score that will be attributed to the overall performance of Black students. This "strength in

numbers" mentality may relieve some of the anxiety associated with stereotype threat.

Of the potential explanations, one of the most recent is a concept known as *stereotype reactivity* (Oswald & Harvey, 2000-2001). In instances of stereotype reactivity, individuals who are the targets of negative stereotypes strengthen their resolves to disprove the negative stereotype with excellent performance in the targeted domain. In the case of the present sample, stereotype reactivity would ensue if respondents were aware of the negative stereotype that African Americans do not perform exceptionally on standardized measures of intelligence, and exerted more effort to ensure that their performance would disprove that stereotype. Essentially, stereotype reactivity should predict differences in performance across stereotype threat conditions that are opposite to those predicted by stereotype threat.

The HBCU setting for this study is likely to engender attitudes that foster stereotype reactivity. As young adults pursuing higher learning and subsequent professional careers, the students at HBCU's represent the upper echelon of the African American community. Because of the collective nature of the African American community, these students may feel that it is their obligation to disprove relevant negative stereotypes.

Limitations

To ensure that stereotype threat was truly activated, it may have been useful to include a comparison sample of non-African American participants to see if there were differences across racial subgroups. There remains the possibility that any participants, regardless of race, could have achieved similar performance across the stereotype threat conditions. Additionally, it may have been useful to test African American students at institutions other than HBCU's in order to determine whether stereotype threat activation is influenced by the presence of racial majority subgroup members. Previous research (e.g. Inzlicht & Ben-Zeev, 2000; 2003; Sekaquaptewa & Thompson, 2002; 2003) supports the proposition that stereotype salience and subsequent threat

35

increases in the presence of majority subgroup members. Therefore, it is likely that Black

students who attend a university with a significant proportion of White students are more likely

to experience stereotype threat than those that attend a more culturally and ethnically

homogenous HBCU.

Practical Implications and Future Research

Although the results of this study contradict those of prior research, there is reason to be

positive. An optimistic interpretation of the given results is that the stereotype concerning

African Americans and subpar academic performance is no longer threatening enough to

negatively affect test performance. To the contrary, the data imply that the stereotype is

audacious enough to influence African American test-takers to work harder to disprove it.

Additionally, research should investigate other potential constructs that may boost performance

(e.g. stereotype reactivity) in the face of negative circumstances. Finally, it is important to better

identify the situations in which stereotype threat is activated. Simply providing instructions that

designate a test as "diagnostic of cognitive ability" or "a set of puzzles" does not appear to

produce discrepant performance.

Conclusion

Although none of the proposed hypotheses were supported by the data, the results give us

hope that widely-held negative perceptions may be slowly dissipating. The recent election of our

country's first African American president has provided a highly visible illustration of the fact

that African Americans are capable of great academic success, and should therefore not be

stereotyped as cognitively inferior. Further, this illustration is likely salient enough to prompt

young African Americans to refute and resist negative stereotypes. Finally, the demonstration

that positive outcomes are possible despite negative expectations is an encouragement to all of those who may fall victim to the effects of negative stereotypes.

References

Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science*, 5, 385-390.

Arthur, W. Jr., & Doverspike, D. (2005). Achieving diversity and reducing discrimination in the workplace through human resource management practices: Implications of research and theory for staffing, training, and rewarding performance. In R. L. Dipboye, & A. Colella (Eds), Discrimination at work: *The psychological and organizational bases.* (pp. 305-327). Mahwah, NJ: LEA.

Arvey, R. D., Strickland, W., Drauden, G., & Martin, C. (1990). Motivational components of test taking. *Personnel Psychology, 43,* 695-716.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173-1182.

Berry, III, R. Q. (2003). Mathematics standards, cultural styles, and learning preferences. The plight and the promise of African American students. *The Clearing House, 76*, 244-249.

Bell, B. S., & Kozlowski, S. W. J. (2002). Goal orientation and ability: Interactive effects on self-efficacy, performance, and knowledge. *Journal of Applied Psychology, 87*, 497-505.

Blanz, M., Mummendey, A., Mielke, R., & Klink, A. (1998). Strategic responses to negative social identity: An empirical systematization of field data. *European Journal of Social Psychology*, *28*, 697–729.

Bobko, P., Roth, P. L., & Potosky, D. (1999). Derivation and implications of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. *Personnel Psychology, 52,* 561-589.

Brown, R. P., & Day, E. A. (2006). The difference isn't Black and White: Stereotype threat and the race gap on Raven's Advanced Progressive Matrices. *Journal of Applied Psychology, 91,* 979-985.

Button, S. B., Mathieu, J. E., & Zajac, D. M. (1996). Goal orientation in organizational research: A conceptual and empirical foundation. *Organizational Behavior and Human Decision Processes, 67*, 26-48.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology, 27*, 703-722.

Ceci. S.J., & Williams, W. M. (1997). Schooling, intelligence, and income. *American Psychologist, 52,* 1051-1058.

Chan, D., & Schmitt, N. (1997). Video-based versus paper-and-pencil method of assessment in situational judgment tests: Subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology, 82*, 143-159.

Crocker, J., Major, B., & Steele, C. M. (1998). Social stigma. In D.T. Gilbert, S.T. Fiske, & G. Lindzey (Eds.), *The Handbook of Social Psychology* (4th edition, Vol. II, pp.504-553). Boston: The McGraw-Hill Companies.

Crocker, J., Luhtanen, R., Blaine, B., & Broadnax, S. (1994). Collective self-esteem and psychological well-being among White, Black, and Asian college students. *Personality and Social Psychology Bulletin, 20*, 506-513.

Cullen, M. J., Hardison, C. M., & Sackett, P. R. (2004). Using SAT-grade and ability-job performance relationships to test predictions derived from stereotype threat theory. *Journal of Applied Psychology, 89*, 220-230.

Davis, C., Aronson, J., Salinas, M. (2006). Shades of threat: Racial identity as a moderator of stereotype threat. *Journal of Black Psychology, 32,* 399-417.

Doojse, B., & Ellemers, N. (1997). Stereotyping under threat: The role of group identification. In R. Spears, P. J. Oakes, N. Ellemers, & S. A. Haslam (Eds.), *The social psychology of stereotyping and group life* (pp. 257-272). Cambridge, MA: Blackwell.

Doverspike, D., Taylor, M. A., & Arthur, W. E., Jr. (1999). *Affirmative action: A psychological perspective.* New York: Nova Science Publishers, Inc.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41*, 1040-1048.

Eden, D. (2002). Replication, meta-analysis, scientific progress, and *AMJ*'s publication policy. *Academy of Management Journal, 45*, 841-846.

Edwards, B. D., & Arthur, W. A. (2007). An examination of factors contributing to a reduction in subgroup differences on a constructed-response paper-and-pencil test of scholastic achievement. *Journal of Applied Psychology, 92,* 794-801.

Elliot, A. J., & McGregor, H. A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80,* 501-519.

Elliot, A. J., & Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: A mediational analysis. *Journal of Personality and Social Psychology, 72*, 218-232.

Fordham, S., & Ogbu, J. U. (1986). Black students' school success: Coping with the burden of acting White. *The Urban Review, 18*, 176-206.

Hartigan, J. A., & Wigdor, A. K. (1989). *Fairness in employment testing: Validity generalization, minority issues, and the general aptitude test battery.* Washington, D.C.: National Academy Press.

Helms, J. E. (1990). *Black and White racial identity: Theory, research, and practice.* New York: Greenwood.

Helms, J. E. (1992). Why is there no study of cultural equivalence in standardized cognitive ability testing? *American Psychologist, 47*, 1083-1101.

Hernstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life.* New York: Free Press.

Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 11,* 365-371.

Inzlicht, M., & Ben-Zeev, T. (2003). Do high-achieving female students underperform in private? The implications of threatening environments on intellectual processing. *Journal of Educational Psychology, 95,* 796-805.

Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review, 5,* 602-619.

Kuncel, N. R., Hezlett, S. A., & Ones, D. (2001). A comprehensive meta-analysis of the predictive validity of the graduate records examinations: Implication for graduate student selection and performance. *Psychological Bulletin, 127,* 162-181.

Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin, 18,* 302-318.

Mayer, D. M., & Hanges, P. J. (2003). Understanding the stereotype threat effect with "culture free" tests: An examination of its mediators and measurement. *Human Performance, 16*, 207-230.

McKay, P., Doverspike, D., Bowen-Hilton, D., & Martin, Q. (2002). Stereotype effects on Raven's Advanced Progressive Matrices scores of African-Americans. *Journal of Applied Social Psychology, 34,* 767-787.

McKay, P. F. (1999). *Stereotype threat and its effect on the cognitive ability test performance of African–Americans: The development of a theoretical model.* Unpublished doctoral dissertation, University of Akron, Akron, Ohio.

McFarland, L. A, Lev-Arey, D. M., & Ziegert, J. C. (2003). An examination of stereotype threat in a motivational context. *Human Performance, 16*, 181-205.

Mummendey, A., Kessler, T., Klink, A., & Mielke, R. (1999). Strategies to cope with negative social identity: Predictions by social identity and relative deprivation theory. *Journal of Personality and Social Psychology, 76*, 229-245.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*, 77-101.

Nguyen, H. D., O'Neal, A., & Ryan, A. M. (2003). Relating test-taking attitudes and skills and stereotype threat effects to the racial gap in cognitive ability test performance. *Human Performance, 16,* 261-293.

Oswald, D. L., & Harvey, R. D. (2000-2001). Hostile environments, stereotype threat, and math performance among undergraduate women. *Current Psychology: Developmental, Learning, Personality, Social, 19,* 338-356.

Oyserman, D., Harrison, K., & Bybee, D. (2001). Can racial identity be promotive of academic efficacy? *International Journal of Behavioral Development, 25,* 379-285.

Ployhart, R. E., Ziegert, J. C., & McFarland, L. A. (2003). Understanding racial differences on cognitive ability tests in selection contexts: An integration of stereotype threat and applicant reactions research. *Human Performance, 16*, 231-259.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales.* Oxford, England: Oxford Psychologists Press.

Reynolds, C., & Brown, R. (1984). *Perspectives on bias in* mental *testing.* New York: Plenum.

Roth P. L., Bevier, C. A., bobko, P., Switzer, F. S., III, & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology, 54*, 297-330.

Rushton, J. P. (2000). *Race, evolution, & behavior: A life history perspective.* (2nd Ed.). Port Huron: Charles Darwin Research Institute.

Rushton, J. P., & Ankey, C. D. (1996). Brain size and cognitive ability: Correlations with age, sex, social class, and race. *Psychonomic Bulletin and Review, 3*, 21-36.

Ryan, A. M., & Greguras, G. J. (1998). Life is not multiple choice: Reactions to the alternatives. In Hakel, M. D. (Ed.). *Beyond multiple-choice: Evaluating alternatives to traditional testing for selection.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Saccuzzo, D. P., & Johnson, N. E. (1995). Traditional psychometric tests and proportionate representation: An intervention and program evaluation study. *Psychological Assessment, 7,* 183-194.

Sackett, P. R., Hardison, C. M., & Cullen, M. J. (2004). On interpreting stereotype threat as accounting for African American-White differences on cognitive tests. *American Psychologist, 59*, 7-13.

Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). High-stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative-action world. *American Psychologist, 56,* 302-318.

Sawyer, T. P., & Hollis-Sawyer, L. A. (2003, April). Stereotype threat and face validity effects on cognitive test performance. Poster session presented at the annual meeting of the Society for Industrial Organizational Psychology (SIOP), Orlando, FL.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124,* 262-273.

Schmitt, N., Clause, C. S., & Pulakos, E. D. (1996). Subgroup differences associated with different measures of some common job relevant constructs. *International Review of Industrial and Organizational Psychology, 11,* 115-139.

Sekaquaptewa, D., & Thompson, M. (2002). The differential effects of solo status on members of high- and low-status groups. *Personality and Social Psychology Bulletin, 28,* 694-707.

Sekaquaptewa, D., & Thompson, M. (2003). Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal of Experimental Social Psychology, 39,* 68-74.

Sellers, R. M., Rowley, S. A. J., Chavous, T. M., Shelton, J. N., & Smith, M. A. (1997). Multidimensional inventory of black identity: A preliminary investigation of reliability and construct validity. *Journal of Personality and Social Psychology, 73,* 805-815.

Smither, J. W., Reilly, R. R., Millsap, R. E., Pearlman, K., & Stoffey, R. W. (1993). Applicant reactions to selection procedures. *Personnel Psychology, 46,* 49-76.

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

Steele, C. M., & Aronson, J. (1998). Stereotype threat and the test performance of academically successful African Americans. In Jencks, C., & Meredith, P. *The Black-White Test Score Gap* (pp. 401-427). Washington D.C.: Brookings Institution.

Steele, C. M., & Aronson, J. (2004). Stereotype threat does not live by Steele and Aronson (1995) alone. *American Psychologist, 59*, 47-48.

Steele, C. M., & Davies, P. G. (2003). Stereotype threat and employment testing: A commentary. *Human Performance, 16,* 311-326.

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype threat and social identity threat. *Advances in Experimental Social Psychology, 34,* 379-440.

Stricker, L. J. (1998). *Inquiring about examinees' ethnicity and sex: effects on AP Calculus AB Examination performance* (College Board Rep. No. 98-1; ETS Research Rep. No. 98-5). New York: college Entrance Examination Board.

Stricker, L. J., & Ward, W. C. (1998). *Inquiring about examinees; ethnicity and sex: Effects on computerized placement tests performance*. (College Board Rep. 98-2; ETS Research Rep. No. 98-9). New York: College Entrance Examination Board.

Stryker, S., & Serpe, R.T. (1982). Commitment, identity salience and role behavior. In W. Ickes & E. Knowles (Eds.), *Personality, roles, and social behavior* (pp. 199-218). New York: Springer-Verlag.

Stryker, S., & Serpe, R.T. (1994). Identity salience and psychological centrality: Equivalent, overlapping, or complementary concepts? *Social Psychology Quarterly, 57,* 16-36.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G.Austin & S.Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S.Worchel & W. G.Austin (Eds.), *Psychology of intergroup relations* (2nd ed., (pp. 7–24). Chicago: Nelson-Hall.

Tucker, L. R., & Lewis, C. (1973). The reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38,*1-10.

Wallace, D. L. (2003). *Anxiety and coping as mediators of stereotype threat and academic performance*. Unpublished master's thesis, California State University, Northridge.

Ziegert, J. C., Ployhart, R. E., & McFarland, L. A. (2002). *Perceived stereotype threat: Development of a self-report scale*. Paper presented at the annual conference of the Society of Industrial and Organizational Psychology, Toronto, Canada.

Zuckerman, M. (1990). Some dubious premises in research and theory on racial differences: Scientific, social, and ethical issues. *American Psychologist, 45*, 1297-1303.
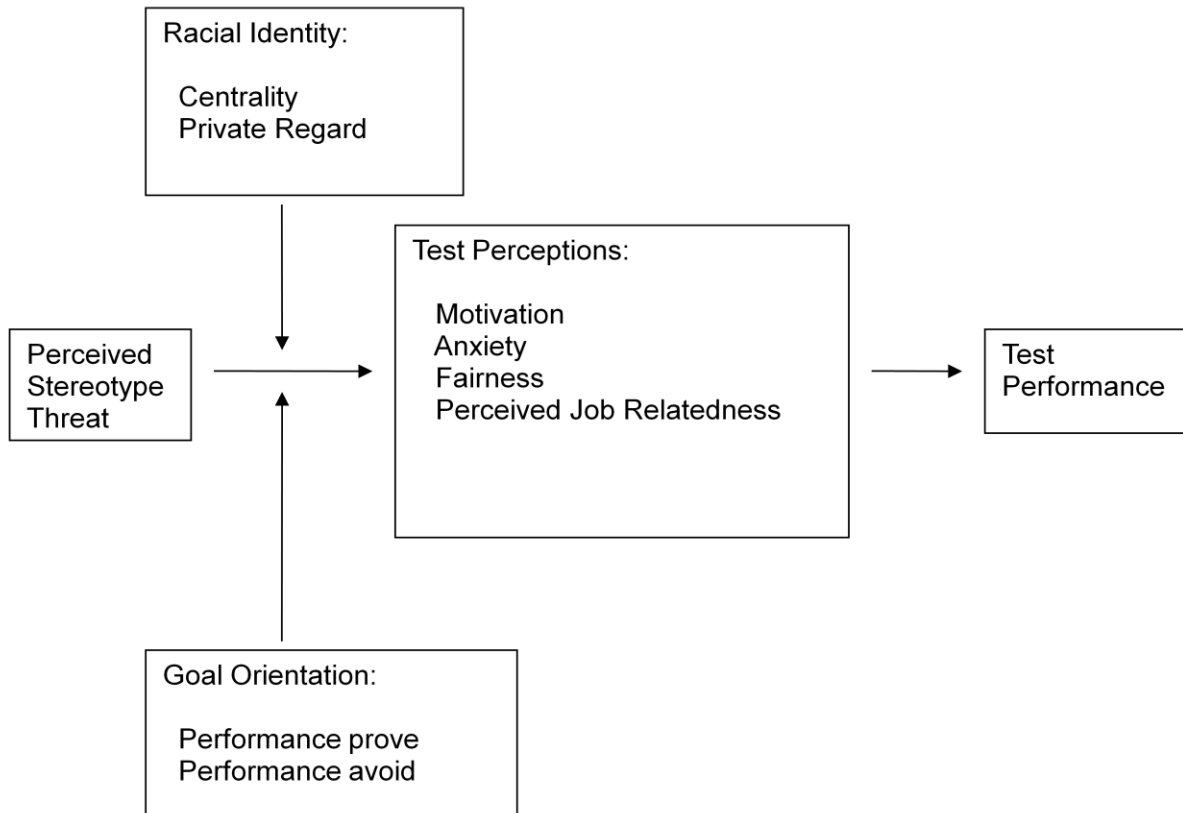
Appendix A

Figure of Stereotype Threat Relationships

Figure 1. Proposed Model of Stereotype Threat Effects

Appendix B

Tables and Questionnaires

Table 1

Means, Standard Deviations, Intercorrelations, for All Dependent Measures (Stereotype Threat Heavy)

Variable	*M*	*SD*	1	2	3	4	5	6	7	8
1. Raven's	4.97	2.09								
2. Job-relatedness	2.60	0.84	.26							
3. Fairness	2.85	0.83	.27*	.66**						
4. Motivation	3.36	0.79	.20	.45**	.53**					
5. Anxiety	2.44	0.77	-.14	.36**	.24*	.32**				
6. GO: Approach	3.57	0.89	-.10	.08	.06	.01	.19			
7. GO: Avoid	3.66	0.99	-.30*	-.11	-.08	.12	.24*	.33**		
8. RI: Centrality	3.86	0.83	-.15	-.10	-.11	-.08	-.09	.24*	.04	
9. RI: Private Regard	4.15	0.80	-.01	-.15	-.07	-.03	-.20	.09	.07	.51**

Note: *p < .05; **p < .01. *n* = 71. Job-relatedness = perceived job-relatedness; Motivation = test-taking motivation; Anxiety = test-taking anxiety; GO: Approach = Goal orientation: performance approach; GO: Avoid = Goal orientation: performance avoid; RI: Centrality = Racial identity: centrality; RI: Private Regard = Racial identity: private regard.

48

Table 2

Means, Standard Deviations, Intercorrelations, for All Dependent Measures (Stereotype Threat Light)

Variable	*M*	*SD*	1	2	3	4	5	6	7	8
1. Raven's	4.56	2.22								
2. Job-relatedness	2.47	0.82	.10							
3. Fairness	2.83	0.80	.34*	.64**						
4. Motivation	3.53	0.72	.46	.12	.35**					
5. Anxiety	2.23	0.74	-.33*	.12	-.02	-.21				
6. GO: Approach	3.40	0.97	.14	.10	.14	-.02	.02			
7. GO: Avoid	3.66	0.86	.06	-.22	-.10	.06	.11	.30*		
8. RI: Centrality	3.89	0.81	.14	-.02	.11	.33*	-.25	.07	.31*	
9. RI: Private Regard	4.33	0.80	.28*	-.12	.13	.21	-.28*	.23	.20	.61**

Note: *p* < .05; **p* < .01. *n* = 57. Job-relatedness = perceived job-relatedness; Motivation = test-taking motivation; Anxiety = test-taking anxiety; GO: Approach = Goal orientation: performance approach; GO: Avoid = Goal orientation: performance avoid; RI: Centrality = Racial identity: centrality; RI: Private Regard = Racial identity: private regard.

Table 3

Means, Standard Deviations, Intercorrelations, for All Dependent Measures (Control)

Variable	*M*	*SD*	1	2	3	4	5	6	7	8
1. Raven's	4.05	1.94								
2. Job-relatedness	2.63	0.95	.20							
3. Fairness	3.05	0.85	.21	.55**						
4. Motivation	3.53	0.85	.35**	.21	.24					
5. Anxiety	2.46	0.83	-.12	.49**	.25	.13				
6. GO: Approach	3.51	0.92	.00	.21	.00	.32*	.23			
7. GO: Avoid	3.52	1.08	-.11	.19	.12	.01	.19	.29		
8. RI: Centrality	3.90	0.71	.01	.07	.08	.30*	.03	.10	.28*	
9. RI: Private Regard	4.08	0.80	.39**	-.18	-.13	.19	-.50**	.05	-.07	.24

Note: *p* < .05; **p* < .01. *n* = 55. Job-relatedness = perceived job-relatedness; Motivation = test-taking motivation; Anxiety = test-taking anxiety; GO: Approach = Goal orientation: performance approach; GO: Avoid = Goal orientation: performance avoid; RI: Centrality = Racial identity: centrality; RI: Private Regard = Racial identity: private regard.

50

Table 4

Tests of Mediating Effects

Mediator	Sobel test statistic	Path a (ST Condition – Test Performance)	Path b (ST Condition – Mediator)	Path c (Mediator – Test Performance)
Job-relatedness	0.12	0.37	0.01	0.45
Fairness	1.23	0.37	0.10	0.69
Motivation	1.20	0.37	0.09	0.86
Anxiety	0.00	0.37	0.00	-0.52

Note: *p* < .05; **p* < .01. *n* = 183. Job-relatedness = perceived job-relatedness; Motivation = test-taking motivation; Anxiety = test-taking anxiety; ST Condition = stereotype threat condition; All paths are betas. Path a = regression of test performance on stereotype threat condition; Path b = regression of mediator (job-relatedness, fairness, motivation, or anxiety) on stereotype threat condition; Path c = regression of test performance on mediator (job-relatedness, fairness, motivation, or anxiety).

Task Perceptions and Beliefs

DIRECTIONS
The following items ask you to rate your general attitudes, beliefs, and opinions about race issues. Read each statement carefully and give your honest feelings about the beliefs and attitudes expressed. Indicate the extent to which you agree with each item by checking or filling in the circle on the accompanying scale that corresponds to your answer.

Strongly Disagree	Disagree	Neither Disagree nor Agree	Agree	Strongly Agree
1	2	3	4	5

Perceived Stereotype Threat	
1. People of my race do significantly better on intelligence tests.	1 2 3 4 5
2. I think others believe that my race determines how well I do on intelligence tests.	1 2 3 4 5
3. I actually have an advantage on intelligence tests due to my race.	1 2 3 4 5
4. I am at a disadvantage on intelligence tests due to my race.	1 2 3 4 5
5. The test may have been easier for people of my race.	1 2 3 4 5
6. The experimenter expected me to do poorly on the test because of my race.	1 2 3 4 5
7. In college classes, people of my race often face biased evaluations.	1 2 3 4 5
8. I never worry that people will draw conclusions about my intelligence based on my race.	1 2 3 4 5
9. Tests, like the one that I just took, have been used to discriminate against people from my race.	1 2 3 4 5
10. During the test, I wanted to show that people of my race could perform well on it.	1 2 3 4 5
11. A negative opinion exists about how people from my race perform on this type of test.	1 2 3 4 5
Perceived Job Relatedness	

12. The actual content of this test is clearly related to the work required in college courses.	1 2 3 4 5
13. I am confident that this test can predict how well a student will perform in college courses.	1 2 3 4 5
14. My performance on this test is a good indicator of my ability to perform well in college courses.	1 2 3 4 5
15. Students who perform well on this type of test are more likely to do well in college courses than students who perform poorly on this type of test.	1 2 3 4 5
16. College admissions counselors could tell a lot about a student's ability to perform in college based on the results of this test.	1 2 3 4 5
Perceived Fairness	
17. The test results will accurately reflect how well I performed on this test.	1 2 3 4 5
18. This test will fairly reflect my ability to perform well in college courses.	1 2 3 4 5
19. Overall, I believe that this test is fair.	1 2 3 4 5
Motivation	
20. Doing well on this test was important to me.	1 2 3 4 5
21. I wanted to do well on this test.	1 2 3 4 5
22. I tried my best on this test.	1 2 3 4 5
23. I tried to do the very best I could do on this test.	1 2 3 4 5
24. While taking this test, I concentrated and tried to do well.	1 2 3 4 5
25. I wanted to be among the top scorers on this test.	1 2 3 4 5
26. I pushed myself to work hard on this test.	1 2 3 4 5
27. I was extremely motivated to do well on this test.	1 2 3 4 5
Anxiety	

28. I probably didn't do as well as most of the other people who took this test.	1 2 3 4 5
29. During a test, I often think about how poorly I am doing.	1 2 3 4 5
30. For this test, I found myself thinking of the consequences of failing.	1 2 3 4 5
31. During a test, I get so nervous I can't do as well as I should have.	1 2 3 4 5
32. I am nervous about how my performance on this test will reflect on my abilities.	1 2 3 4 5
33. I am nervous about how my performance on this test will compare to that of others who take it.	1 2 3 4 5
34. I am nervous about what performance on this test means in terms of my intellectual capabilities.	1 2 3 4 5
35. I am nervous about the test scores reflecting my best effort.	1 2 3 4 5
36. I am nervous about the test scores reflecting my true potential for success.	1 2 3 4 5
37. I am nervous about doing my own personal best on this test.	1 2 3 4 5

Goal Orientation

38. In class, it is important for me to perform better than other students.	1 2 3 4 5
39. In class, it is important for me to do well compared to others.	1 2 3 4 5
40. My goal in class is to get a better grade than most of the students.	1 2 3 4 5
41. In class, I just want to avoid doing poorly.	1 2 3 4 5
42. My goal in class is to avoid performing poorly.	1 2 3 4 5
43. My fear of performing poorly in class is often what motivates me.	1 2 3 4 5

MIBI

44. I feel good about Black people.	1 2 3 4 5
45. In general, being Black is an important part of my self-image.	1 2 3 4 5

46. I am happy that I am Black.	1 2 3 4 5
47. I have a strong sense of belonging to Black people.	1 2 3 4 5
48. I often regret that I am Black.	1 2 3 4 5
49. Being Black is an important reflection of who I am.	1 2 3 4 5
50. Blacks contribute less to society than others.	1 2 3 4 5
51. Overall, I often feel that Blacks are not worthwhile.	1 2 3 4 5