11-6-2017

I'm Afraid to Tell You What I Really Think: An Investigation into the Feedback Withholding Bias Mechanism and Outcomes within STEM Settings

Deborah Lee
drlwt9@mail.umsl.edu

I'm Afraid to Tell You What I Really Think: An Investigation into the Feedback Withholding Bias Mechanism and Outcomes within STEM Settings

Deborah R. Lee
M.A., Industrial/Organizational Psychology, University of Missouri-St. Louis, 2012
M.A., Industrial/Organizational Psychology, Western Kentucky University, 2009
B.A., Psychology and Women's Studies, Saint Louis University, 2007

A Dissertation Submitted to the Graduate School at the University of Missouri-St. Louis in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Psychology with an emphasis in Industrial/Organizational Psychology

December 2017

Advisory Committee

Stephanie Merritt, Ph.D.
Chairperson

Bettina Casad, Ph.D.

William Kyle, Ph.D.

Matthew Taylor, Ph.D.

Abstract

This study builds on previous research findings that White individuals who desire to not appear racist is associated with Black students failing to receive constructive feedback, compared to White students (Croft & Schmader, 2012). This Feedback Withholding Bias (FWB) may inhibit the ability for Black students to learn from constructive feedback which is important for student learning and future performance. Black male students and White male evaluators with a STEM major were the focus of this study because of the underrepresentation of Black STEM students and workers and previous research focusing on racism stereotypes impact on the FWB. The results suggest that stereotype threat is not the underlying mechanism of the FWB. Findings do suggest impression management plays a role in the FWB as participants' motivation to control prejudice (MCP) was predictive of both stereotype threat and the FWB measures. Results demonstrated the White evaluators provided similar amounts of constructive feedback to Black and White students, yet also showed a positive feedback bias for Black students.

I'm Afraid to Tell You What I Really Think:
An Investigation into the Feedback Withholding Bias Mechanism and Outcomes within STEM Settings

Although women and minorities have a greater presence in the workforce today than in previous decades, both groups remain underrepresented in the science, technology, engineering, and math (STEM) fields (National Science Foundation, 2011). According to the National Science Foundation (2011), Blacks and Latinos comprise 13% of the total workforce and only 3% of the technical or STEM jobs. The number of historically disadvantaged minorities remains low in both the national STEM workforce and the college-level STEM degrees. Research has found that women and minorities are less likely to initially choose a STEM major and even if they do, are less likely to remain in the major (Chen & Thomas, 2009). This research investigates feedback given to minority- individuals as one source for the discrepancies seen in career paths in the science and technology fields.

Feedback is often an integral part of performance management, but it is also important to the learning process. Constructive feedback is respectful, specific, and provides external attributions for poor performance, while providing internal attributions for good performance (Baron, 1988). Often, however, actual feedback is given in vague and subjective terms, such as "great job" or "could be better." One factor that can influence the type of feedback that is given is domain relevant stereotypes (Biernat & Danahar, 2012).

The Feedback Withholding Bias (FWB) occurs when minority students fail to receive critical feedback from evaluators, possibly because evaluators are

concerned about appearing prejudiced (Croft & Schmader, 2012). This is problematic in that individuals need to receive accurate and constructive feedback in order to monitor and evaluate their own work and improve subsequent performance. Therefore, the feedback withholding bias may inhibit minority improvement, performance, and retention. The goal of this study is to investigate the process that underlies this phenomenon of withholding critical feedback.

The purpose of this research is to investigate potential causes of the FWB. The focus is on feedback given to individuals in a domain in which their social group is negatively stereotyped. I refer to Black individuals in STEM settings as "counter-stereotypical" because they are in contrast to the non-Hispanic White and Asian men that are dominant in this domain thus, the STEM domain is stereotypically White or Asian and male (Beede, Julian, Khan, Lehrman, McKittrick, & Doms, 2011; Landivar, 2013).

The goal of this research is to identify means to encourage evaluators to provide constructive feedback, including constructive criticism, to minorities in counter-stereotypical settings. When feedback is provided in a constructive manner, this can lead to improvements in performance (Kluger & DeNisi, 1996). This may increase retention of these counter-stereotypical group members in underrepresented areas, in this case the STEM disciplines. Although extensions to other settings are beyond the scope of this project, understanding and eliminating the FWB may also have similar effects on retention rates of underrepresented groups in other settings, such as work organizations.

Importance of Constructive Feedback

Kluger and DeNisi (1996) defined feedback as "actions taken by (an) external agent(s) to provide information regarding some aspect(s) of one's task performance" (p. 255). More specifically, Baron's (1988) study distinguished between constructive and destructive types of feedback and investigated the impact of feedback on students' self-efficacy for future tasks. As previously stated, Baron defined *constructive criticism* as specific, considerate, and focused on the actions involved in task performance, rather than on the individual's personality or other stable characteristics. A violation of at least one of the basic components for delivering effective constructive feedback may have deleterious consequences (Baron, 1990). In contrast to constructive criticism, *destructive feedback* includes only general performance comments, is inconsiderate in tone, and poor performance is attributed to the individual and may be accompanied by direct or implied threats. It does not necessarily follow that if feedback is not constructive then it is destructive; it may be that the feedback is neutral because it is neither constructive nor destructive.

Most applicable to the current research is constructive criticism. Research suggests that critical, constructive feedback is vital for individuals to adequately evaluate their own work (Jussim & Eccles, 1992). In addition, constructive criticism is crucial to learning environments that promote skill acquisition and content-knowledge (Jussim & Eccles, 1992). Within educational settings, constructive criticism has been shown to be critical to a learning environment that promotes

students' content knowledge and skill development (Price, Handley, Miller, & O'Donovan, 2010).

Constructive criticism includes both positive and negative critical feedback (Cole, 2008) and offers useful and diverse strategies for skill improvement that are specific to the individual's performance level. Constructive criticism also provides encouragement and support (Bjorklund, Parente, & Sathianathan, 2004). In summation, constructive feedback is specific to performance level or task, incorporates both positive and negative feedback, offers specific strategies for improvement, and done in an encouraging and supporting manner.

Research demonstrates that constructive feedback positively impacts participants' self-efficacy (Martocchio & Dulebohn, 1994), while destructive feedback negatively impacts participants' self-efficacy on future tasks (Baron, 1988). Cole (2008) found a positive relationship between Black and Hispanic minority students receiving constructive criticism from faculty and greater academic skills, performance, and educational satisfaction. Also supporting the advantages of constructive, a meta-analysis by Kluger and DeNisi (1996) found that feedback interventions fail to be effective when feedback providers move away from task-related feedback to focus more on self-related issues. That is, the authors concluded that task-focused, constructive feedback is more likely to help the feedback receiver. However, research suggests that constructive feedback may be disproportionately withheld from minority students (See Croft & Schmader, 2012). The current study examines the underlying mechanism impacting the failure of supervisors to provide vital, constructive feedback to individuals who are stereotyped as less competent in

the STEM domain. Specifically, the differences exhibited during feedback between individuals from different racial backgrounds.

Feedback Withholding Bias

As previously stated, the FWB occurs when minority group members fail to be provided with constructive criticism on their performance. Several studies have established that work purportedly created by minority individuals tends to receive inflated ratings compared with identical work purportedly created by White individuals (Harber, 1998; 2004, Harber, Stafford, & Kennedy, 2010). However, it was unclear whether that inflation was due to an overemphasis of positive aspects of the work or to withholding feedback on negative aspects of the work. Furthermore, alternative processes could have explained the inflation of scores. For example, the inflation may have been caused by a shifting standards effect (Biernat & Manis, 1994; Biernat & Danahar, 2012) wherein minority students' work was held to a lower standard than the work of others. Croft and Schmader (2012) posited that evaluators might believe that negative feedback would demotivate individuals from a stereotypically disadvantaged educational background and therefore evaluators protect those students by providing praise rather than criticism. Alternative explanations include patronizing or self-focused fears of appearing prejudiced (Croft & Schmader, 2012). That is, it is possible that evaluators may be reluctant to provide criticism for fear that any criticism will be perceived as racially biased (Cohen & Steele, 2002).

In a seminal study, Croft and Schmader (2012) addressed these issues, effectively identifying the FWB. The results of Croft and Schamder's (2012) study

indicate that evaluators can be concerned about appearing prejudiced when they provide constructive feedback to minority students, and that those concerns were significantly associated with FWB effects. Specifically, Croft and Schmader (2012) tested the hypothesis that evaluators' individual differences in motivation to control their prejudice would be associated with FWB and found that individuals who were motivated to not be perceived as prejudiced by others were more likely to exhibit the FWB. Australian participants were tasked with highlighting sections of good writing in one color and bad writing in another color on an essay the participants believed to either be written by a White or a minority (Aborigine) student. The results of two experiments demonstrated that both essays received around the same amount of positive feedback, but minority students received less negative feedback than White students, indicating that differences in feedback provided do seem to result from withholding of constructive information on performance aspects that need improvement.

In addition, the study distinguished between internal and external motivation to control prejudice. People high in internal motivation to control prejudice have internalized egalitarian goals, whereas those high in external motivation to control prejudice wanted to avoid the social sanctions associated with appearing prejudice. Across both experiments, the FWB was most evident for participants high in external motivation to control prejudice (Croft & Schmader, 2012). The strong prediction of the FWB by the motivation to control prejudice construct supports the argument that the FWB has greater impression management motives. Croft and Schmader also found no support for the alternative hypothesis

that providing biased feedback is related to evaluators' patronizing attempts to protect minority students' academic engagement from the sting of criticism. Instead, their study supports the impact of evaluators' impression management concerns and the desire to avoid appearing racist.

Also supporting the notion that impression management is a key factor in FWB, a few recent studies have found that White evaluators provide more useful feedback to minority students if the rater's egalitarian values were affirmed prior to the feedback (Harber et al., 2010; Ruscher, Wallace, Walker, & Bell, 2010). In the Haber et al. (2010) study, participants received an egalitarian boost by completing a survey designed to reinforce positive minority views. The opinion-based survey included items such as "Government offices should be closed on Martin Luther King Day" and "The confederate flag should not fly over government buildings," and the response scale forced respondents to agree at least somewhat to these statements. Theoretically and empirically demonstrated, this reinforcement buffers against the threat of being perceived as prejudiced, allowing evaluators to provide more honest and constructive feedback without feeling threatened (Bergsieker, Shelton, & Richeson, 2010). Previous research has supported that impression management may serve as a form of self-preservation for individuals experiencing stereotype threat (von Hippel, von Hippel, Conway, Preacher, Schooler, & Radvansky, 2005), such that it allows for a barrier between the stereotype threat situation and the individual's behavior. Specifically, von Hippel and colleagues (2005) found that individuals high in impression management were likely to cope with stereotype threat by either denying incompetence in the threatened domain (e.g., Black deny

cognitive incompetence when interviewed by White person) or denying the domain importance (e.g., White deny intelligence is important when threatened by stereotype of being less intelligent than Asians). Thus, these studies support the assertion there may be a stereotype threat effect impacting majority evaluators' feedback to counter-stereotypical individuals.

Harber and colleagues (2010) investigated positive feedback bias (where evaluators provide more praise and less criticism for minorities, same as FWB but focused on positive feedback rather than lack of feedback) as a response to self-image threats in teacher trainees tasked with evaluating students. The authors of the study manipulated self-image threats using two versions of a survey. One version encouraged participants to respond in either a pro-minority, egalitarian manner while the other was designed to encourage responses unfavorable to minorities. Responding to the latter survey form threatened participants' egalitarian self-images because it forced them to endorse less-egalitarian responses, such as "People should be allowed to fly the confederate flag on their own front lawn". The phrasing of the questions (i.e., "should be") elicits responses that are palatable on the individual level, but collectively represent an anti-minority sentiment. Results indicated the evaluators whose egalitarian self-images were threatened by this manipulation provided Black students with more favorable essay-content ratings, recommended less time for writing skills development, gave more positive copy-editing comments, and tied those comments to equivocating buffers. One the other hand, evaluators whose egalitarian values were first affirmed failed to demonstrate a FWB with Black students. Their feedback was

indistinguishable from feedback supplied to White essay writers. Haber and colleagues suggest, "for Whites, as for minorities, intergroup contact can present self-image risks that deter the authentic, non-self-conscious communications through which true egalitarianism is displayed" (p. 217).

Prejudice in STEM Disciplines

A study conducted in 2011 using US census data reports that Blacks and Latinos continue to be underrepresented in STEM employment (Landiver, 2013). Thus, it is likely that historically disadvantaged minorities in the United States STEM fields will often receive performance-related feedback from a member of a racial out-group (e.g., a White or Asian professor or employer). Previous research suggests that learning from an individual within a student's own identity group (e.g., race, gender) can have a positive impact on performance and intentions to stay within the content field (See Robst, Keil, & Russo, 1998). In addition, research from secondary education found that having more Black instructors increased the chances of Black students enrolling in more advanced math classes (Klopfenstein, 2005) and that having a same-race teacher was associated with higher reading and math scores for Black students (Dee, 2005). It may be that minority recipients discount feedback from out-group individuals or that the feedback provided is lacking. This study focuses on cross-racial feedback not just because the feedback may lack features of constructive criticism, but also because out-group (i.e., White and Asian men) evaluators are more likely to be the individuals providing the feedback in general.

Previous research shows that the potential for prejudice, even discrimination, is present when perceivers hold stereotypes about a particular

social group (e.g., minorities) and when the stereotypes are incongruent with the attributes that they believe are required for success in a particular role (Eagly & Karau, 2002; Heilman, 2001). A stereotype is a commonly held belief about a social group or type of individual, and stereotypes typically become part of the society's shared knowledge. Stereotypes, like other heuristics, are cognitive shortcuts that can be helpful in daily life, but they can also easily lead to a negative bias (Eagly & Carli, 2007). The category of race is one of the first things that people notice about others, and implicit racial stereotypes relevant to the situation are activated automatically (Eagly & Carli, 2007). These associations between race and the stereotyped characteristics and qualities of the minority group are pervasive and even unconsciously influential (Dovidio, Kawakami, & Gaertner, 2000; Sczesny & Stahlberg, 2002). Stereotypes remain pervasive because of their perceived utility in the accurate prediction of the future, including predictable interactions between individuals that are perceived to represent different stereotypes (Mendes, Blascovich, Hunter, Lickel, & Jost, 2007). Indeed, regardless of whether a minority individual actually exhibits stereotypical characteristics, people's subjective beliefs about the characteristics of minority groups may lead them to believe that any given individual group member lacks the qualities to be successful in a counter-stereotypical domain (e.g., a Black scientist; Eagly & Chin, 2010).

Empirical research supports the premise that stereotypes impact our perceptions of others (Dovidio et al., 2000), our behaviors towards others (Eagly & Chin, 2010), and our judgments about others (Eagly & Karau, 2002; Heilman, 1995). Several areas of research have focused on detrimental effects that occur when a

stereotyped group's attributes are incongruent with the stereotyped attributes required for success in a particular role (Dipboye, 1985; Eagly & Karau, 2002; Heilman, 1995). That is, when someone belongs to a group that is stereotyped to lack the characteristics believed to be necessary for success in a particular role, the individual will likely receive less favorable role-related judgments from others (Diekman & Hirnisey, 2007; Eagly & Chin, 2010; Eagly & Karau, 2002). In addition, research shows interacting with someone that violates stereotype-based expectations will disrupt one's ability to accurately predict how the interaction will unfold and can create uncertainty with how to behave (Mendes et al., 2007).

Stereotypes related to STEM disciplines may be influenced by media depictions of scientists, engineers, mathematicians, and the like. Past studies show that media impacts the stereotypes people endorse (Devine, 1989; Potts & Martinez, 1994; Tyree, 2011). A content analysis of stereotypes in science education television reported that minorities were significantly less likely to be labeled as a 'scientist' and received significantly less on-screen time compared to White characters (Long, Boarsky, & Thayer, 2001). In addition, other research supports that television media tends to reinforce negative cultural stereotypes for minority groups (Bell-Jordan, 2008; Tyree, 2011).

Supporting research demonstrated that across many western cultures the scientist prototype is typified as a "mature, intelligent, hardworking, White male, wearing glasses, formally dressed or in a lab coat" (Walls, 2012, p. 15). Stereotypes regarding Black male Americans typically include descriptions of criminality, laziness, lack of intelligence, and unpredictability (Bogle, 2001). Thus, Black

individuals in the United States' STEM fields may face challenges due the conflicts between negative stereotypes of Blacks and the characteristics of successful scientists. I argue that the influence of stereotypes extends to differences in the feedback provided to individuals – specifically, the FWB.

The key focus of this research is on feedback provided to Black male STEM students from White or Asian evaluators. I hypothesize that FWB will be seen in such situations for several reasons. Because Black Americans are stereotyped as having an educational disadvantage and stereotyped as less intelligent than White or Asian people and lazy (Fiske 1998; Reyna, 2000) and White people tend to be stereotyped as racist (Chang & Kleiner, 2003), I expect that White evaluators aware of these prominent stereotypes will likely display the FWB because they are motivated to be perceived as egalitarian.

As aforementioned, past research has consistently demonstrated that White individuals are aware that they may be stereotyped as prejudiced or racists (Dunton & Fazio, 1997; Goff, Steele, & Davies, 2008; Plant & Devine, 1998). Studies have found that the unintended and ironic consequence of being concerned about appearing prejudiced actually tends to increase the likelihood that they will be perceived as prejudiced. In fact, one study demonstrated that individuals concerned with appearing prejudiced physically distanced themselves from Black partners to a greater extent, thus appearing more prejudicial than individuals not concerned about appearing prejudiced (Goff et al., 2008).

I propose that in a context where individuals are required to provide feedback to a Black STEM student, participants will be more likely to buffer

themselves from the stereotype of being prejudiced or racist. The premise of this study is that this buffer will be a lack of constructive, critical feedback (i.e., FWB). That is, in the context of providing feedback to Black students majoring in the STEM filed will elicit the FWB in participants.

FWB Measurement and Processes

Measurement of the FWB

This study adds to the existing literature by employing multiple methods to measure the dependent FWB variable. The most common measurement method in previous research used evaluators trained to code different aspects of the experimental stimulus, typically an essay of moderate quality purportedly written by a student. In the majority of these studies, trained evaluators coded the positive and negative comments provided by participants, which may include comments related to essay content, mechanics, general comments and buffer comments (i.e., used to qualify other comments; Harber et al., 1998; 2010). In general, past research utilized subjective self-report ratings from evaluators on the feedback provided by participants (Harber 1998, Harber et al., 2004, 2010, 2012; Jeffries & Hornsey, 2012; Ruscher et al., 2010). One previous study employed a highlighting method, where participants were asked to highlight positive and negative essay sections with different colored highlighters and a comparison of the total length of the different highlight colors across conditions served as a measure of the FWB (Croft & Schmader, 2012). That is, participant's highlighting less in the color denoting 'needs improvement' compared to the color denoting 'positive feedback' on the stimulus essay evidenced the FWB.

Similar to the previous methods, this study also asks participants to rate the stimulus essays on various components (e.g., overall essay quality, quality of interaction with confederate) as well as a more objective measurement of the FWB by using independent raters to code the positive and negative words written by each participant. In addition, a third measure of the FWB employed in this study was the difference between feedback notes that the participant believes are their 'personal notes' and the verbal feedback provided to the confederate will be measured. The use of three different measurement methods for the FWB and the congruence, or lack thereof, between the three will be a contribution to the FWB literature.

Researchers have found that compared to when evaluators believe their feedback will remain anonymous, when they believe their identity will be known to the minority student receiving feedback, evaluators offer more praise and less constructive criticism to minority students than White students (Harber, 1998, 2004; Harber et al., 2010). This is a type of FWB itself, in that evaluators withhold negative feedback for minority student, yet still provide positive feedback only when personally identifiable. Previous studies utilized a between-persons design to examine the FWB. One of the methods of measuring the FWB in this study is designed to examine this type of FWB within-evaluator (i.e., within-person design). That is, measuring the difference of written feedback to verbally delivered feedback should provide a strong and valid indicator of the FWB in the present study. It is likely that the difference in feedback occurs because the evaluators are attempting

to manage recipient perceptions, specifically perceptions that the evaluator is racist (Croft & Schmader, 2012).

Stereotype threat. Stereotype threat is a widely studied topic and research confirms that the phenomenon occurs when individuals of a negatively stereotyped group experience concern or anxiety in a situation where they have the potential to confirm that negative stereotype about their social group (Steele, 1997, Schmader, Johns, & Forbes, 2008). If a negative stereotype is present regarding a specific social group, members of that group are likely to become anxious about their performance, which may impair performance. Specifically, much research has demonstrated the performance decrements on intellectual ability diagnostic tests for Blacks (Steele & Aronson, 1995), Latinos (Gonzales, Blanton, & Williams, 2002), and children of lower socioeconomic status (Croizet & Claire, 1998). In addition, evidence suggests that under stereotype threat psychology students perform worse than science students on test of intellectual ability (Croizet, Despres, Gauzins, Huguet, Leyens, & Meot, 2004) and women perform more poorly than men on math tasks (Spencer, Steele, & Quinn, 1999). Even social groups that typically are at an advantage can experience stereotype threat. In particular, when told their performance would be compared to the performance of Asian men, White men performed poorer on a math test (Aronson et al., 1999), and White participants performed less well on a motor task than Black participants when told the test was assessing natural athletic ability (Stone, 2002; Stone, Lynch, Sjomeling, & Darley, 1999).

Croft and Schmader (2012) argue that the results of their study on the FWB demonstrate that the impression management concerns of an evaluator greatly impact given feedback. The authors, however, inferred that stereotype threat was impacting feedback, rather than directly testing this hypothesis. Past findings show stereotype threat impacts Whites when they are afraid of confirming the stereotype that Whites are racist (Frantz, Cuddy, Burnett, Ray, & Hart, 2004). For instance, Frantz and colleagues found greater evidence of stereotype threat for participants that displayed external motivation to control prejudice as opposed to internal motivation to control prejudice. Thus, some White evaluators – those high in external motivation to control prejudice—may experience stereotype threat when faced with providing negative feedback to students of color.

In a stereotype threat situation, there should be three factors present that are related to each other: the concept of one's in-group (e.g., I am White), the concept of the domain in question (e.g., stereotypically Whites are racist), and the individual experiencing stereotype threat's self-concept (e.g., I am White; Whites are racist; but I am not racist; Schmader et al., 2008). Stereotype threat occurs during situations that make negative stereotypes about one's own group salient; thus, that situation poses a threat to the individual's self-integrity, which is their sense of their own value (Steele, 1997). The threat to self-integrity is a result of the ambiguity between the individual's success expectations and the negative social stereotype implying poor performance. According to Schmader et al., (2008), the self-integrity threat acts as a stressor and prompts individuals to monitor their behaviors and cognitions, and it also impacts their physiological stress response, affective

reactions, and efforts to cope with the aversive experience. All of these reactions typically lead to the individual failing to optimally perform. Neuroscience studies show that when their racial prejudices may come to light, White individuals are more vigilant in monitoring their internal signs of bias (Amodio, et al., 2004; Amodio, Kubota, Harmon-Jones, & Devine, 2006). Research has also shown that individuals under stereotype threat consciously monitor their performance, with more attention paid to threat cues as well as cues that they are effectively coping with the situation. Thus, it can be expected that White evaluators who need to provide feedback to minority students may activate mental monitoring and vigilance processes, which is one key aspect of stereotype threat.

Assessment of Stereotype Threat Processes. Schmader and colleagues (2008) presented an integrated process model of stereotype threat effects that suggests stereotype threat disrupts performance through three distinct, yet interrelated mechanisms (see Figure 1). These mechanisms include a physiological stress response that directly impairs prefrontal processing, a tendency for individuals to actively monitor their performance, and efforts to suppress negative thoughts and emotion for the purpose of self-regulation.

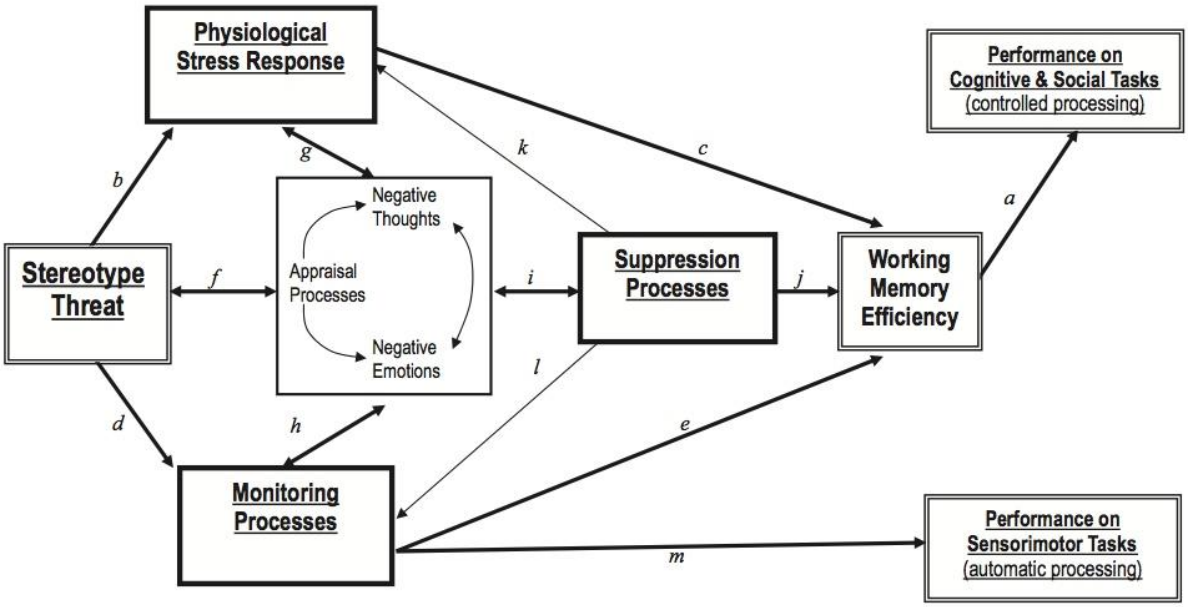

Figure 1. Schmader et al.'s (2009) Integrative process model for stereotype threat. Adapted from Schmader, T., Johns, M., & Forbes, C. (2008). An integrated model of stereotype threat effects on performance. *Psychological Review, 115,* 336-356. doi: 10.1037/0033-295X.115.2.336

Schmader and colleagues' integrated process model assumes that stereotype threat negatively impacts performance on social and cognitive tasks that require controlled processing (Schmader et al., 2008). I contend that providing thorough feedback, especially face-to-face feedback, requires social and cognitive resources and can be considered a controlled process that may be disrupted by stereotype threat. Specifically, stereotype threat interferes with cognitive and social tasks through taxation on working memory by several different processes, such as the physiological stress reaction and cognitive monitoring, which lead to thought-suppression. Research on stress and cognition suggest that when individuals are in stereotype threat situations, they experience stress-induced physiological arousal as evidenced by increased blood pressure for Black students compared to White

students told they were taking an intellectual diagnostic test (Blascovic, Spencer, Quinn, & Steele, 2001), and increased sympathetic nervous system (SNS) activity by women watching an imbalanced gender group discussing math and science (Murphy, Steele, & Gross, 2007). Likewise, White individuals, threatened by the Whites-are-racist stereotype, demonstrated a cardiovascular response related to a physiological threat profile (i.e., decreased cardiac output combined with increases in total peripheral resistance) when interacting with a Black male, compared to when interacting with a White male (Mendes, Blascovich, Lickel, & Hunter, 2002). Mendes and colleagues also found that Whites performed poorly on a verbal task in the Black male interaction condition, thus providing general evidence for poorer cognitive performance.

In addition to the increase in SNS activation during an acute physiological stress response, the integrated stress response will also include the release of corticosteroids and catecholamines (Schommer, Hellhammer, & Krischbaum, 2003), which may also impair cognitive performance (Eysenck & Calvo, 1992). Furthermore, when cognitive processing relies on the hippocampus or the prefrontal cortex, there may be a greater impact because of the high concentration of cortisol receptors in those areas of the brain (Metcalfe & Jacobs, 1998). Schmader and colleagues (2008) argue that research supports that acute social stressors increase cortisol levels and possibly directly reduce the efficiency of executive processing. While giving constructive feedback, executive processing should be significantly impacted for individuals under stereotype threat because they have to

create and give feedback at the same time they are consciously monitoring their performance.

The second process that Schmader and colleagues (2008) suggest operates during stereotype threat is active and conscious cognitive monitoring of performance. That is, to avoid failure in a stereotype threat situation, people switch from their more automated state of functioning to a more conscious and controlled state of self-monitoring during the threatening situation. In a series of experiments based on regulatory focus theory (Higgins, 1997, 1999), results showed that individuals under stereotype threat become more focused on avoiding failure (i.e., prevention focus), resulting in more systematic and cautious performance (Siebt & Forster, 2004). Individuals with greater prevention focus demonstrate avoidance strategies and are more sensitive to evidence of negative performance outcomes (Higgins, 1999). Along the same lines, Beilock, Rydell, and McConnell (2007) found that women under math stereotype threat reported worrying about and monitoring their performance. The authors suggest that these thoughts overload working memory thereby contributing to the reduction in performance on more difficult tasks. Thus, stereotype threat effects on performance are more likely to be present in cognitively demanding situations where the difficult task(s) and stereotype threat process jointly have depleted cognitive resources, rather than on easy tasks where cognitive resources remain plentiful even under stereotype threat.

Additional studies found that women were motivated to disprove the negative stereotype of females as “bad drivers” when under stereotype threat; however, even motivated drivers under stereotype threat demonstrated decreased

driving performance during a driving simulation compared to women not induced with stereotype threat (Yeung & von Hippel, 2008). Furthermore, Yeung and von Hippel demonstrated that individuals performing a divided attention task performed at the same level as individuals under stereotype threat, lending more support to the notion that stereotype threat impacts working memory capacity. Likewise, Schmader and colleagues (2008) argue that in addition to monitoring for signs of failure, individuals under stereotype threat will likely also monitor for internal cues signaling that one is successfully coping with the stressful situation. It may also be possible that by suppressing stereotype-related thoughts, evaluators actually experience fewer negative thoughts about counter-stereotypical individuals' performance; however, the feedback they are providing lacks constructive criticism because the thought process was negatively impacted by stereotype threat.

Thought-suppression processes tax working memory resources chronologically following the physiological stress response and cognitively monitoring the stereotype threat situation. Specifically, research shows that individuals attempt to suppress negative thoughts and emotions when in stereotype threat situations (Schmader et al., 2008). Although thought-suppression fails to consistently be demonstrated through self-report measures (see Wheeler & Petty, 2001), studies utilizing less-conscious indicators have shown support for the thought-suppression process. For example, although homosexual men in a stereotype threat condition of interacting with a preschool child did not report increased anxiety, they exhibited more nonverbal anxiety than heterosexuals and

gay men in non-threatening conditions (Bosson, Haymovitz, & Pinel, 2004). Schmader and colleagues (2008) posit, and empirical research supports that thought suppression (see Muraven & Buameister, 2000) and emotion suppression (see Gross, 2002) are generally effortful and energy consuming and thus impair working memory. In addition, suppressing thoughts and emotions tends to be associated with an ironic processing effect, such that those same thoughts and emotions are more easily accessible for some time after the suppression effort ends (Wegner, Erber, & Zanakos, 1993). Therefore, self-regulation of thoughts and emotions should have a negative impact on performance on complex tasks in stereotype threat situations.

This study directly tests whether stereotype threat is one of the mechanisms that contribute to the FWB, where evaluators fail to provide critical, constructive feedback to individuals in counter-stereotypical settings. Specifically, I propose that when evaluators are in a situation where they could appear prejudiced, stereotype threat is activated.

Hypothesis 1: White evaluators will experience greater levels of stereotype threat when assigned to provide feedback to a Black student than when assigned to provide feedback to a White student.

Stereotype threat may cause the evaluator to spend precious cognitive resources on controlling their perceived prejudiced behavior, rather the task of providing ample, critical feedback to the counter-stereotypical individual, because the evaluator's fears being perceived as prejudiced.

Hypothesis 2: Evaluators will demonstrate the FWB to a greater extent with Black STEM students compared to White STEM students.

Hypothesis 3: Greater stereotype threat will be positively associated to the degree of FWB.

Individual Difference Moderators of FWB

Motivation to control prejudice. In order to more directly test the notion that stereotype threat/impression management concerns play a role in FWB, Croft and Schmader (2012) examined internal and external motivation to control prejudice, two distinct yet related constructs (i.e., not opposites on a continuum). As society's norms have changed regarding racism in the last half of the 20th century, research has consistently shown a decrease in self-reported overt racial prejudice (e.g., Dovidio & Gaertner, 2000). Due to the strong societal pressures to appear nonprejudiced in the United States today, behavior that appears non-prejudiced may be associated with compliance with norms rather than internalized egalitarian motives (Butz & Plant, 2009). Thus, Plant and Devine (1998, 2009) posit that both an internal, personal motivation to respond without prejudice and a more normative, external motivation to respond without prejudice exist. Internal motivation to control prejudice (Internal MCP) is the intention to be free of prejudiced thoughts and emotions, while external motivation to control prejudice (External MCP) is the intention to hide prejudice by controlling behavior (Plant & Devine, 2009). An individual may be internally motivated to respond without prejudice because of an internalized egalitarian goal or externally motivated to respond without prejudice to conform to social norms and avoid others' disapproval

(Dunton & Fazio, 1997; Plant & Devine, 1998, 2009). That is, Plant and Devine argue that some individuals strive to be truly free of prejudicial attitudes or behavior and are internally motivated to respond without prejudice. On the other hand, conforming to societal pressure to avoid expressing prejudice represents an external motivation to respond without prejudice. Individuals with a higher Internal MCP tend to control their prejudice across situations, regardless of external pressures, whereas individuals who are externally motivated seek to disguise their prejudice publically when external pressures arise, but continue to privately endorse negative biases (Butz & Plant, 2009; Dunton & Fazio, 1997). Plant and Devine (1998) assert that individuals with high External MCP may lash out in a more subtle, socially acceptable way against minorities because they feel that their personal freedoms are constrained (e.g., such as being told to suppress a stereotype, See Monteith, Spicer, & Tooman, 1998)

Research has empirically shown that the motivation to control or to respond without prejudice is an effective predictor of individual behavior in counter-stereotypical contexts, such as interracial interactions (Glaser & Knowles, 2008; Plant & Devine, 1998, 2009). As stated before, the presence of normative pressure impacts the likelihood of an individual responding with prejudice. It is also likely that externally motivated individuals are more sensitive to social pressure than others (Butz & Plant, 2009). Plant and Devine (1998) found that for individuals primarily externally motivated, their prejudicial reactions vary depending on the context, in particular whether it's a private or public context. Externally motivated participants asked to indicate the degree of endorsement for Black stereotypes

strongly endorsed them anonymously; however, they reported low prejudice beliefs when publicly telling their answer to an experimenter. In contrast, participants who were internally motivated, regardless of external motivation level, reported low prejudicial beliefs across both the public and private conditions.

Researchers investigating the impact of internal and external motivation to respond without prejudice on interracial interactions found differential outcomes for the related concepts (e.g., Plant, 2004). Results from Plant, Devine, and Peruche's (2010) study indicated that individuals highly internally motivated tended to use behaviors and strategies that approach a positive and egalitarian outcome, which resulted in a longer, smoother, and generally better-rated interaction by the participant, their Black partner, and objective coders. In comparison, participants highly externally motivated tended to use strategies and behaviors that would avoid negative and prejudicial outcomes and were ironically rated as more prejudiced by their Black partner.

In line with this research, Croft and Schmader (2012) found that Internal and External MCP had different associations with the FWB. The authors reported that individuals with a high External MCP but low Internal MCP were more likely to inhibit negative feedback to counter-stereotypical individuals. Specially, they found that individuals who were externally motivated to appear nonbiased, but lacked internal motivation to control prejudice, were more likely to show FWB in terms of holding back negative feedback and inflating minority grades. These individuals may have, internally, evaluated minority performance more negatively due to the influence of negative stereotypes, but they may have then suppressed their external

feedback regarding those negative evaluations. On the other hand, individuals reporting low levels of both Internal MCP and External MCP to appear unbiased provided significantly more negative feedback to minority students than to White students, indicating those unconcerned about conforming to current social norms may be more susceptible to stereotypes and perhaps overly harsh when evaluating minority students' work. Importantly, individuals reporting high levels of Internal MCP showed no significant FWB effect, suggesting that Internal MCP may be a key moderating factor in the FWB process.

It is likely that highly externally motivated individuals are more sensitive to societal pressures to behave in an egalitarian manner (Butz & Plant, 2009) and thus when placed in a situation where they may confirm the widely-known societal stereotype that Whites are racist and that they do not behave in an egalitarian manner, those individuals are more likely to experience stereotype threat. As previously stated, individuals with a high External MCP and low Internal MCP were more likely to inhibit negative feedback to minority individuals (Croft & Schmader, 2012). I propose a moderated mediation model where Internal MCP moderates the relationship between External MCP and stereotype threat as well as stereotype threat mediates the relationship between MCP and FWB, such that those with higher levels of External MCP and lower levels of Internal MCP will experience stereotype threat to a greater extent and thus exhibit greater levels of the FWB. According to Muller, Judd, and Yzerbyt (2005), a moderated mediation model occurs when the mediating process responsible for producing the overall effect depends on the value of the moderator.

Hypothesis 4: Internal MCP will moderate the relationship between External MCP with the FWB such that the FWB will be highest when External MCP is high and Internal MCP is low

Hypothesis 5: Stereotype threat will mediate the association between the interaction of Internal and External MCP and the FWB.

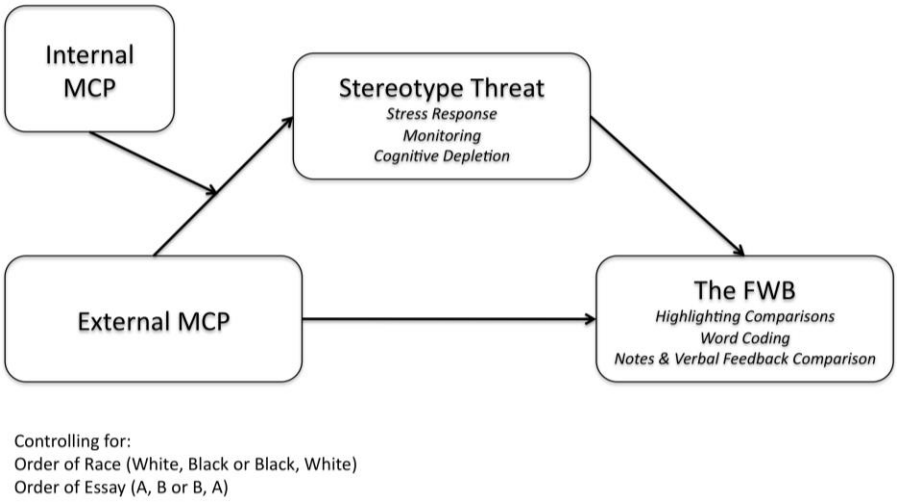

Figure 2. Hypothesized model to be tested in Hypotheses 4 and 5.

Similar to how stereotypes can become activated automatically, resulting in behaviors occurring outside of awareness, goals and motivations can also operate outside conscious awareness (Moskowitz, Gollwitzer, Wasel, & Schaal, 1999; Glaser & Knowles, 2008). One study by Moskowitz and Li (2011) found that when egalitarian goals are cognitively accessible (e.g., primed), the presence of an African American individual would trigger the egalitarian goal and inhibit automatic stereotype activation, without conscious awareness. Based on Plant and Devine's (1998) work, Glaser and Knowles (2008) established a similar construct known as the Implicit Motivation to Control Prejudice (Implicit MCP), which was hypothesized

to reflect processes not under conscious awareness and to relate to the inhibition of automatic prejudicial behaviors. The authors argue that an implicit negative attitude towards prejudice and an implicit belief that oneself is prejudiced combined demonstrates individuals' Implicit MCP levels.

Implicit Motivation to Control Prejudice (Implicit MCP). Motivation to Control Prejudice (MCP) examines explicit (i.e., self-report) measures of Internal and External MCP, and will also investigate the predictive value added with the implicitly measured version of MCP. Implicit MCP is measured using a response-time based method that may demonstrate the capability of distinguishing which participants are more likely to exhibit the FWB. Recent research indicates that, similar to attitudes, goals and motives can exist and function outside of conscious awareness and control (Chartrand & Bargh, 1996). Building on nonconscious motivation, research demonstrates that the goal of egalitarian behavior might operate outside of conscious awareness and reduces unintentional, automatic prejudiced attitudes and behavior (Bargh, 1999; Glaser & Knowles, 2008). As previously mentioned, the original focus of the MCP construct was on the relationship of explicit motivations and explicit, overt discriminatory behavior. Past studies have found Plant and Devine's (1998) self-report measures of Internal MCP and External MCP predicted which individuals were more likely to express prejudice. For example, those high Internal MCP but low External MCP demonstrated less racial bias in a sequential priming task as well as the relationship between bias and Internal MCP moderated the overt discriminatory behavior (Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002). That is, individuals who

possess a strong belief system opposing discrimination are more likely to control discriminatory behavior, even if they are influenced by prominent societal stereotypes. More recent research investigates the impact of implicit motivations on more uncontrollable or unintentional behaviors (e.g., Glaser & Knowles, 2008).

To investigate the Implicit MCP construct, Glaser and Knowles (2008) administered the Implicit MCP measures as well as both an IAT measuring implicit attitudes regarding Black people and weapons and a modified shooting task, intended to be a measure of unintentional behavior. Participants were shown photos of either Black or White targets holding either weapons or harmless objects (e.g., soda cans, cell phones). When the targets were armed, participants needed to press a key indicating, "shoot." When the targets were unarmed, they needed to press a different key indicating, "hold fire." Responses were made in a window from 300-1000 milliseconds, and should therefore reflect automatic processes. Overall, a shooter bias effect was found where individuals that demonstrate a strong implicit stereotype connecting Blacks and weapon possession were faster to shoot armed Blacks than armed Whites and slower to indicate safety to unarmed Blacks than unarmed Whites. The results indicate that individuals with negative implicit attitudes towards prejudice demonstrate the weakest relationship between their implicit stereotypes and automatic behavior, whereas those who find prejudice more acceptable tend to demonstrate a stronger relationship between implicit attitudes and behavior (Glaser & Knowles, 2008). Furthermore, only individuals who score high in belief that oneself is prejudiced as well as high in negative

attitudes toward prejudice failed to exhibit a relationship between racial bias (attitudes) and the shooter task (behavior).

Park, Glaser, and Knowles (2008) further demonstrated the weak relationship between implicit racial bias and the shooter task for high Implicit MCP individual, even under cognitive load. Thus, those participants, despite the additional cognitive drain, were more accurate in their performance than low implicit MCP participants whose behavior was more impacted by their implicit stereotypes. Thus, it is likely that the evaluators under stereotype threat and with low Implicit MCP levels will exhibit stereotype threat and the FWB to a greater degree than evaluators with high Implicit MCP.

Hypothesis 6: Implicit MCP will be positively related to stereotype threat.

Hypothesis 7: Stereotype threat will mediate the associate between Implicit MCP and the FWB.

Method

Participants

White Male STEM Majors (N = 69) were drawn from a student population in the Midwest with ages ranging from 18 to 38 (M = 19.90) through the psychology department subject pool. The sample consisted of 26 Freshmen, 22 Sophomores, 11 Juniors, and 9 Seniors. The study required participants be White men and enrolled in a STEM field major and any participants not fitting those requirements were screened out before completing the online survey. All participants that completed

both parts of the study were entered into a random drawing to win a $50 Amazon gift card.

Design and Procedure

The basic study design consists of male, STEM students reviewing and providing feedback on two essays purportedly written by aspiring STEM students at a different university. The feedback was given via video recording under the cover story that the video will be shown to the essay writer at a later time. This study is a within-person experimental design where each participant participated in an experimental stereotype threat condition (giving feedback to a Black student) and a control condition (giving feedback to a White student). Order of the student's race and the essay were counterbalanced across participants. Prior to giving feedback, participants first completed the pre-testing online questionnaire that included basic demographics and individual difference measures, such as attitudes towards feedback and stereotypes of STEM students.

Participants scheduled a lab session and upon arrival were escorted into a private experimental room and told that the study examines academic interactions between senior and more junior STEM students. The experimenter informed participants that the study looks at feedback provided to freshman-level students. The experimenter explained that physiological sensors are used in the experiment and then applied to the participant to take a 5-minute baseline.

After the physiological baseline assessment, participants were read instructions informing them that they will be reviewing two UMSL undergraduate STEM students' application essay drafts for a STEM scholarship. The stimulus

materials included the scholarship application blank with the applicant's name, class standing (Freshman), date of birth, race, sex, declared major, and the application essay. The two application essays were pilot tested to ensure they are comparable in quality. The first names of the stimulus applicants were derived using census data indicating the most popular names by race over a two-year period late 2000s. In the experimental condition, the applicant was a Black freshman named Jayden and in the control condition, the applicant was a White freshman named Alexander.

Participants received a counterbalanced presentation with the different candidates' information sequentially so they do not have both sets of information at the same time. Instructions asked participants to take 5-10 minutes to review the first candidate's essay. During this time participants highlighted the positive and negative segments of writing on the actual essay and made suggestions, corrections, comments within an online survey page marked "Evaluator's Notes – DO NOT Give to Candidates" to boost feelings of anonymity in the notes section. Participants were told they would have the highlighted essay and feedback notes available during the video-recorded feedback session. The experimenter returned to the private experiment room after 10 minutes (or when the participant informed the experimenter, they were ready to provide recorded feedback). Participants signed a bogus consent form agreeing to be taped and agreeing that the video can be shared with the essay authors. This was based on Croft and Schmader's (2012) study and designed to create a strong belief that participants would be directly communicating their feedback to the essay author.

In the final phase of the experiment, the experimenter told the participant that it was time for the participant to provide feedback on the application materials, and the experimenter activated the camera. The feedback conversation as recorded via a computer laptop camera. After the feedback delivery was complete, the participant completed the Stroop Task, followed by two self-report measures one relating to their feedback experience, including their stress level, and rating if the student should receive the scholarship. Finally, the first quick manipulation check was conducted by asking the participant to identify the race of the applicant from a list of options.

The process was repeated with the second essay. Participants received the other applicant's information packet and had 10 minutes to review the information and make notes. The experimenter proceeded with the second video recorded feedback session after which the Stroop Task was completed, the second set of self-report measures, second manipulation check administered and then sensors were removed. Participants were probed for suspicion and debriefed on the study. See Figure 3 for the experimental procedure.

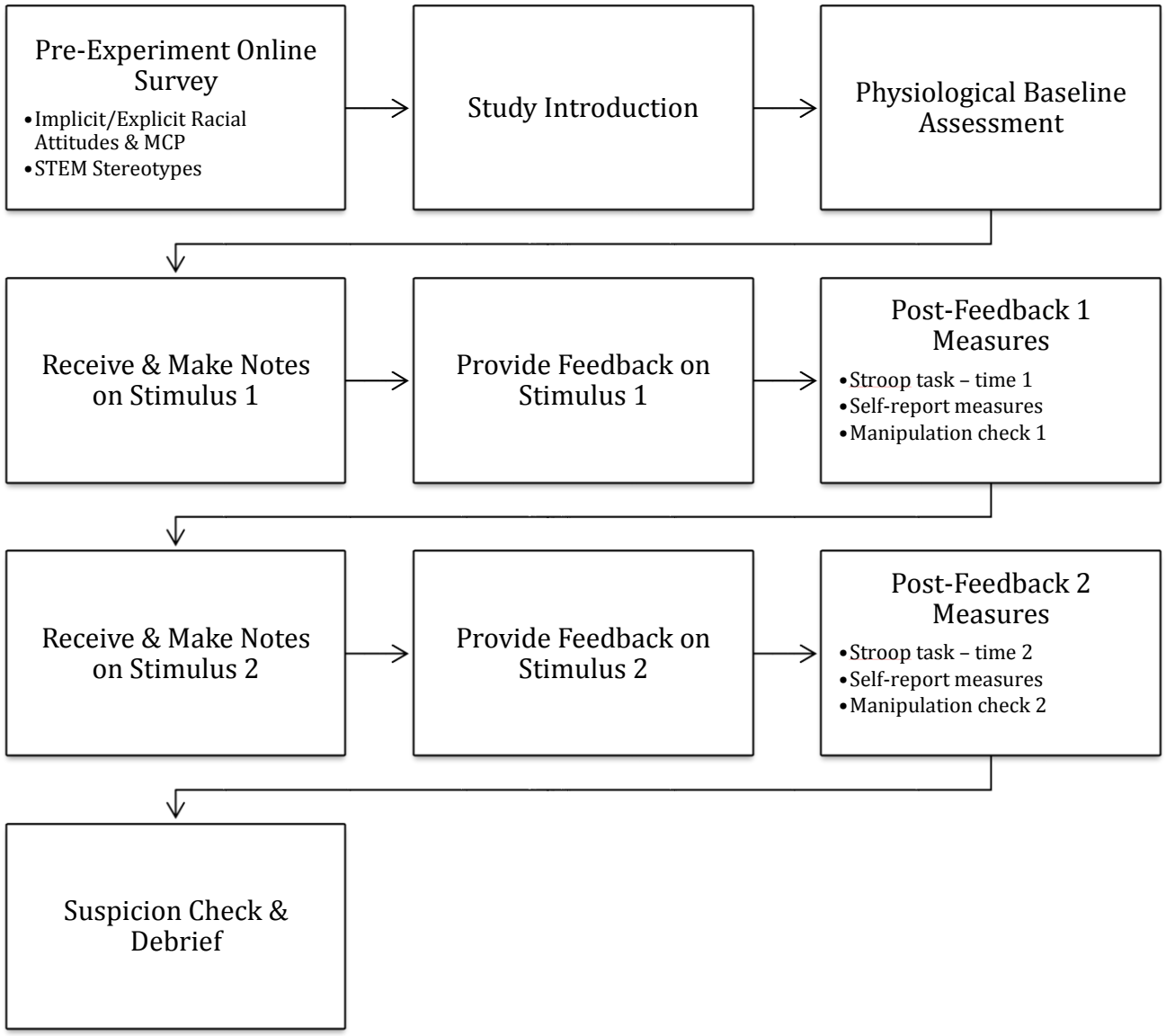

Figure 3. Experimental procedure flow chart.

Materials

Stimuli and Dependent Measures. Two essays of comparable quality were the primary study stimuli. The author worked with a tutor in the Student Writing Center to create two essays that were of 'average' quality and both essays were pilot tested. In the pilot test, undergraduate psychology students highlighted the essay in different color highlighter for both positive and negative content respectively and were asked to provide both a letter grade for the essay and a scholarship recommendation rating. The essay grades were converted to number scores for analysis from 1 (*F*) to 12 (*A+*), with 5 representing a "C" grade. The recommendation scores were on a 1 (*Do Not Recommend*) to 5 (*Highly Recommend*) Likert-type scale. Descriptive statistics and one-ways ANOVAs did not support any significant differences between the two essays indicating that the two stimuli essays are comparable (see Table 1).

Table 1.

Pilot testing descriptives for the two essays used as stimuli

	Essay A (*n* = 18)		Essay B (*n* = 15)	
	M	*(SD)*	*M*	*(SD)*
Positive Highlighting	27.95	15.08	22.08	17.42
Negative Highlighting	32.83	16.13	26.11	14.15
Essay Grade	6.16	1.76	6.36	1.60
Scholarship Recommendation	2.61	.85	3.00	.39

The FWB. The FWB was measured in multiple ways.

Highlighting FWB measure. The first FWB measurement follows the procedure outlined by Croft and Schmader (2012) and was operationalized as the total number of inches of negative and positive highlighting on each essay. The difference in negative feedback for White students compare to Black students will serve as a FWB indicator. One graduate research assistant scanned the essays and used Adobe Acrobat's measuring tool to measure the highlighting in inches and summed all of the positive and negative highlight measurements for total positive and negative essay highlights scores.

Written to spoken comments difference FWB measure. However, the highlighting procedure may reflect only a subset of feedback interactions, thus another measure that entails measuring the difference in the written notes that a participant "privately" recorded compared to the verbally recorded feedback was another dependent FWB measure. In order to code the verbal and written comments, two undergraduate research assistants, blind to study purpose, independently counted the number of positive, negative, constructive, and neutral comments (see Appendix A for more details). Intraclass Correlation Coefficients (ICC) were used to assess inter-rater reliability which suggests acceptable levels of agreement between the raters for both verbal comments (ICCs range from .59 to .94) and written comments (ICCs range from .73 to .94). The degree of correspondence between the spoken and the written feedback provided is an original measurement of the FWB. This indicator was operationalized by subtracting the number of verbal comments (i.e., positive, negative, and constructive

comments) in the recorded feedback from the number of written comments in the purported anonymous notes section.

Post-Feedback Measures. To measure *perceptions of the feedback experience*, participants will complete a short questionnaire regarding their experience in the preceding feedback delivery. Items include "How well do you think your feedback will be received" on a 1 *(not well at all)* -7 *(very well)* Likert-type scale and if they were "Are you concerned about how you will be perceived" on a 1 *(not at all)* – 7 *(extremely)* Likert-type scale as well as given question blank to explain concerns. Reliability analyses did not support combining these two items into a single measure, thus they remain as single-item measures.

Global impressions. In line with Croft and Schmader's (2012) design, each participant will be asked to grade the essay on a scale of 0 – 100. Finally, they will make two general ratings on how qualified they believe the STEM student is and if the student should receive the scholarship on a 1 – 7 Likert-type scale (α=.86).

Independent Measures. In addition to the measures presented below, there were several others measures utilized in the study to allow for possible exploratory analyses. Please see Appendix B for details on the additional measures not discussed here.

Motivation to Control Prejudice (Pre-Experiment Online Survey). Explicit MCP was measured using two self-report scales, Internal Motivation to Control Prejudice (Internal MCP) and External Motivation to Control Prejudice (External MCP; Plant & Devine, 1998). Both the Internal MCP scale (α = .89) and External MCP scale (α = .87) are comprised of five items rated on a 1 (*strongly disagree)* to 5

(*strongly agree)* Likert-type scale. An example Internal MCP item is "Being non-prejudiced is important to my self-concept" and an example External MCP item is "I try to hide any negative thoughts about Black people in order to avoid negative reactions from others".

Implicit Motivation to Control Prejudice (Pre-Experiment Online Survey). Implicit MCP was measured with two different IATs. The first assesses negative attitudes toward prejudice (NAP) by pairing "prejudice" and "tolerant" words with "good" and "bad" words. The second IAT assesses the belief that oneself is prejudiced (BOP) construct regarding ones beliefs that they are prejudice by pairing "prejudice" and "tolerant" words with "me" and "others" (Glaser & Knowles, 2008). In addition, a Race IAT was also included as an additional implicit measure.

The procedure and analysis of the IAT follows the recommendations of Greenwald, Poehlman, Uhlmann, & Banaji (2009). The IAT requires rapid sorting of exemplars (words or pictures) representing two concept categories (e.g., *Prejudice, Tolerant*) and two attribute categories (e.g., *good* or *bad*) into their superordinate categories with a standard set of response seven response blocks. To acclimate the participant to the task, they complete 20 trials sorting *good* or *bad* using two response keys (e.g., *good* words with the *e* key, *bad* words with the *i* key) and then 20 trials sorting the two target concepts with the same two keys (e.g., *Prejudice* words with the *e* key, *Tolerant* words with the *i* key). The next blocks test the participants' implicit associations by sorting items from all four categories with the same two keys alternating by trial between concept and evaluative items (e.g., *Prejudice* and *good* with the *e* key, *Tolerant* and *bad* with *i* key). Essentially, the

difference in average latency between the first sorting conditions (e.g., *Prejudice* and *good*) and the second (e.g., *Tolerant* and *bad*) was taken as the relative association strengths between the concepts and evaluations. That is, participants who find it easier to sort White with good (and Black with bad) are said to implicitly prefer Blacks to Whites. The resulting IAT *D* score is conceptually similar to a Cohen's *d* effect size measure indicating the direction and strength between the concepts and evaluations. The standard method to assess IAT reliability is to correlated the D scores from the first and second blocks. The correlation coefficients for both NAP (.22) and BOP (.20) were lower. The Correlation Coefficient for the Race IAT was stronger (.56) and more consistent with past research findings.

***Stereotype Threat Process Indicators*.** Several different materials were used to measure stereotype threat, including physiological measures, measures of working memory, and self-report measure.

Physiological measures. Cardiac and hemodynamic measures were used to record noninvasively according to commercial safety standards and followed the guidelines established by the Society for Psychophysiological Research (e.g., Sherwood, et al., 1990). This study measured pulse (heart) rate and blood pressure to assess the physiological components of stereotype threat.

The hemodynamic data was collected using a portable blood pressure monitor with a pulse oximeter that estimated blood pressure responses over a fixed time (i.e., every 5 minutes). Blood pressure and heart rate were measured using CMS CONTEC PM50 Ambulatory Blood Pressure Monitor with an attached finger

pulse oximeter respectively. An interactive software program was used to record and score the physiological data of each participant.

The sensory equipment provided several cardiac and hemodynamic physiological indicators for use in data analysis, including pulse rate and mean arterial blood pressure (i.e., average blood pressure, MAP). To analyze the data, mean values were calculated for the cardiovascular indices during each experimental condition and reactivity scores were created by subtracting each participant's minute-5 baseline from each measure during feedback preparation and delivery and then averaging the resulting scores (see Sawyer, et al., 2012).

Cognitive measures. The Stroop task assessed participants' working memory function during the experiment. In this task, participants are shown color words and are asked to name the color in which the words are printed. The meaning of the word and the font color may be congruent or incongruent. For example, the word "blue" can be printed in blue (congruent) or in red (incongruent). If the word and font color are incongruent, participants are more likely to err. That is, participants are more likely to choose the word rather than the ink color in which in the word is printed (Engle, 2002). Performance on the Stroop task requires that the goal of choosing the font color is maintained even when the word elicits a stronger response to say the word. Performance on this task therefore reflects availability of cognitive resources, such that a higher score indicates fewer available cognitive resources.

Self-reported stereotype threat was measured once using four items modeled off previous research, where two items are general stereotype threat items and two

specific stereotype threat items (see Chung, Ehrhart, Ehrhart, Hattrup, & Solamon, 2010; Mayer & Hanges, 2003). The items (α = .83) include (a) When interacting with other races, I worry that people will draw conclusions about my racial group based on my performance; (b) I often think about issues concerning race; (c) I often feel that people's evaluation of my behavior is based on the racial group that I belong to; (d) when interacting with other races, I worry that people will draw conclusions about me based on what they think about my racial group.

Self-reported state anxiety and threat emotions were measured by a short battery modified from Osborne's (2001) study and administered immediately after feedback. Items were rated on a 5-point Likert-type scale (α = .93) from 1 (*very little*) to 5 (*very much*) with the item stem of "to what extent did you feel the following while giving feedback" and the items include tense, under pressure, under strain, nervous/jittery, uneasy, calm, afraid of not doing well, and uncomfortable.

Results

The data were checked for univariate and multivariate outliers, and no cases warranted removal based on analyses. One participant, however, was removed due to failing the manipulation check to correctly identify the student's race. See Table 22 in Appendix C for correlations between main study variables separated by experimental and control condition, Table 23 for descriptives on independent variables and Table 24 for descriptives on outcome variables.

To test Hypothesis 1, that White evaluators would experience stereotype threat when providing verbal feedback to Black students, four repeated-measures ANCOVAs were used because the four stereotype threat measures were not

correlated enough to warrant a MANCOVA analysis. There were two control variables included in the analysis, the order of essays and the order of student race. To control over the overall alpha level or Type I error rate a Bonferroni correction was used setting alpha at .01 for significance (Meyers, Gamst, & Guarino, 2006). Evaluation of the properties of the data set (e.g., normality, equality of variances-covariance matrices) determined that these data did not meet sphericity assumptions and therefore the Greenhouse-Geisser statistic is reported.

There was a statistically significant effect of experimental condition on stereotype threat emotions (F (1, 66) = 17.66, $p < .01$, η^2=.21), where anxiety emotions were slightly higher for the Black student condition, even as the means by condition were similar. Experimental condition also had a significant effect on working memory (F (1, 66) = 17.44, $p < .01$, η^2=.21) where incongruent latencies of the Stroop were higher in the Black student condition, indicating more working memory depletion in the Black student condition. The effect of experimental condition on the physiological MAP measure was significant at the $p < .05$ level, but was not significant after the Bonferroni correction (F (1, 66) = 6.01, p = .02, η^2=.09). Likewise, experimental condition did not have a significant effect on pulse rate (F (1,66) = 2.44, p = .12, η^2=.04). Thus, results suggest only self-reported anxiety and working memory depletion were higher when participants were providing feedback to a Black student versus a White student.

Table 2

Mean, Standard Deviation and Cohen's d of the stereotype threat indicators by experimental condition.

Stereotype Threat Indicator	Experimental Condition	*M*	*SD*	Cohens *d*
Anxiety Emotions	Black Student	1.92	.72	
	White Student	1.91	.69	0.01
Stroop Latency (*ms*)	Black Student	1075.91	336.68	
	White Student	1065.57	391.40	0.03
MAP	Black Student	10.25	23.53	
	White Student	1.41	18.79	0.40
Pulse Rate	Black Student	1.32	13.96	
	White Student	1.57	12.68	0.02

However, results show a significant interaction between the control covariate Race Order and Experimental Condition in all four of the ANCOVA analyses indicating that race order (i.e., Black student first or White student first) was significantly related to stereotype threat. Further mixed ANCOVA tests with Race Order as a between-subjects IV were used to further investigate the impact on stereotype threat measures. With the exception of pulse rate, results show stereotype threat was higher when participants provided feedback to the Black student first.

Results show a significant difference in MAP only when the Black student was provided feedback first. In addition, with Race Order as an IV, working memory differences by student race become nonsignificant (F (1, 66) = .05, p = .82, η^2=.00) which indicates that the original significant results was due to greater working

memory depletion following the first feedback session, regardless of student race, than the second feedback session. Anxiety emotions, however, remained significantly different (F (1, 66) = 5.91, p = .01, η^2=.08) where participants expressed more anxiety emotions after providing feedback to the Black student.

Therefore, Hypothesis 1 was partially supported in that self-reported anxiety was higher when participants were providing feedback to a Black student versus a White student. Interaction plots showing these order effects are found in Figures 4-7. However, these order effects should be interpreted with caution as comparisons with a Bonferroni correction did not reach statistical significance.

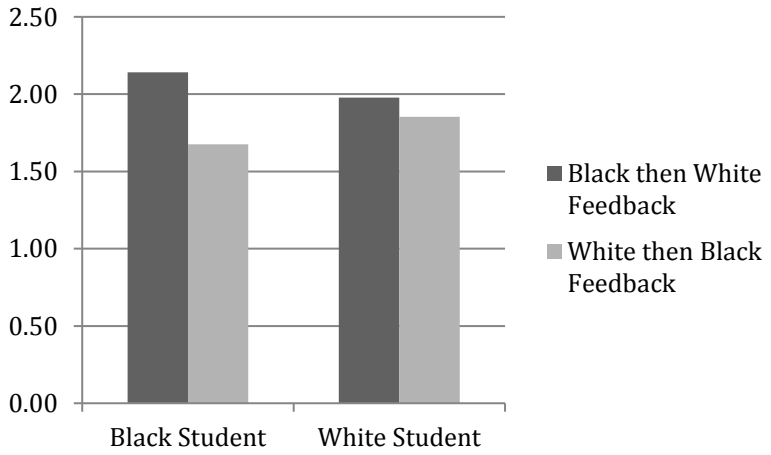

Figure 4. Interaction of participant's experimental condition and the order of student race with the stereotype threat indicator of self-reported anxiety emotions.

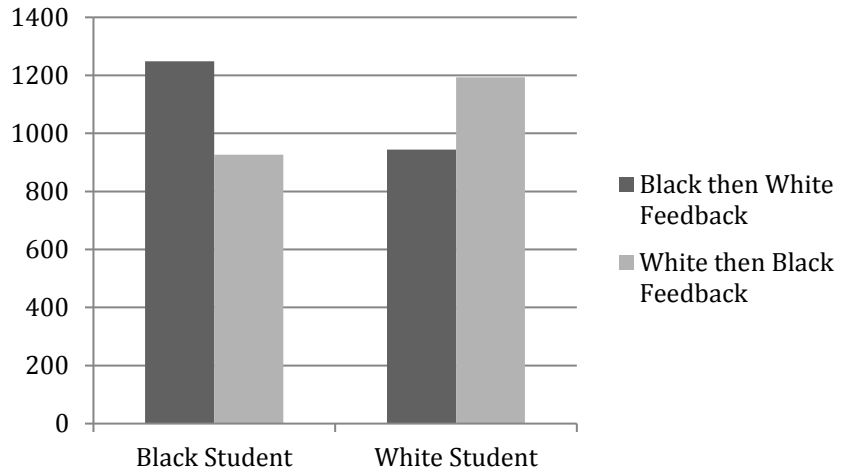

Figure 5. Interaction of participant's experimental condition and the order of student race with the Stroop test stereotype threat indicator.

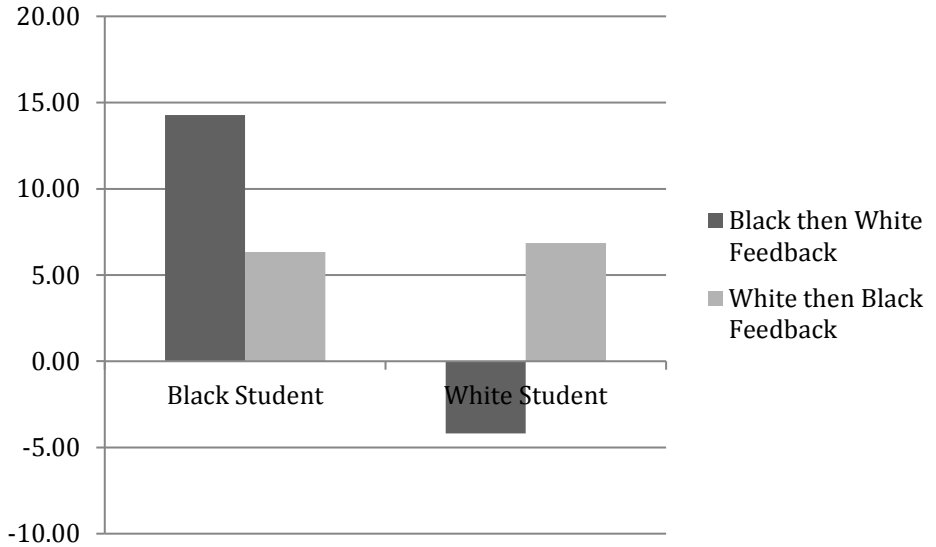

Figure 6. Interaction of participant's experimental condition and order of student race with the physiological stereotype threat indicator MAP.

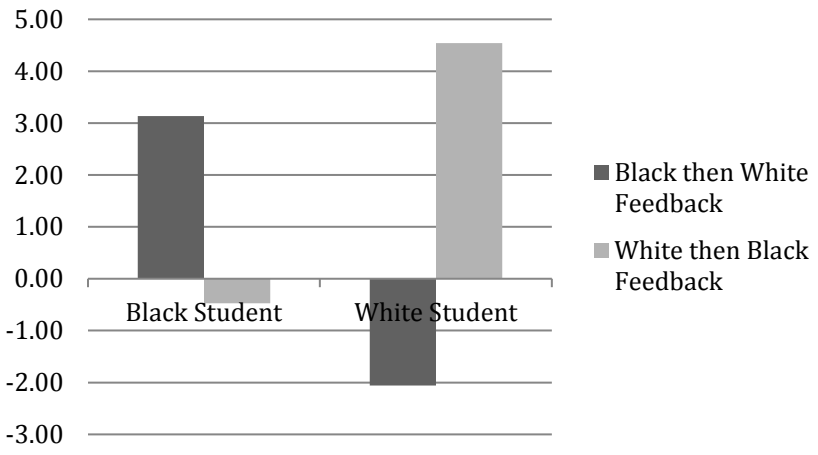

Figure 7. Interaction of participant's experimental condition and order of student race with the physiological stereotype threat indicator pulse rate.

Hypothesis 2 posited that Black students would be given feedback that is more positive and less constructive than feedback provided to White students. As a reminder, the FWB was operationalized in several ways, and means for each of these are displayed in Table 3. Because the dependent variables relevant for Hypothesis 2 were not correlated at a level to warrant a MANCOVA, paired samples *t*-tests were used with a Bonferroni correction applied where the critical value was $\alpha < .01$.

Positive highlighting was examined first, and a significant difference was found -- the Black student's essay ($M = 24.02$, $SD = 16.62$) had more positive areas highlighted than the White student's essay ($M = 19.74$, $SD = 15.61$; $t(66) = 3.17$, $p = .002$). Participants were told the highlighting was private and would not be provided to the essay author, suggesting that this difference may have been due to internal motivation to be unbiased rather than to externally motivated desires to

appear unbiased to the feedback recipient. No significant differences, however, were found in the amounts of negative highlighting.

We also examined the difference between the feedback in participants' written notes (which they were told would private and not be shown to the student) and the verbal feedback provided to the student. Feedback withholding would be evident if participants had more negative (Negative FWB) or constructive (Constructive FWB) comments in their private notes than they provided verbally to the students. Feedback withholding *bias* would be evident if participants withheld feedback at differential rates depending on student race. The difference in feedback withholding between experimental conditions was not statistically significant, failing to support Hypothesis 2. Note that the Positive FWB and Constructive FWB have negative means which indicates participants' providing more positive and constructive feedback verbally than provided in the participant's written notes. Results indicate that participants tended to speak more positive and constructive comments than write them for both races. Although results show this tendency even more so when providing feedback to the Black student, the difference between the conditions was not statistically significant.

Overall, the results of Hypothesis 2 indicate that participants generally provided more positive feedback to the Black student than the White student, but did not generally withhold constructive feedback from the Black student compared to the White student.

Table 3

Dependent Samples T-Test for Difference in FWB Between Black Student and White Student Feedback Conditions

	Black		White				
FWB Indicator	*M*	*SD*	*M*	*SD*	*t*	*df*	Cohen's *d*
Positive Essay Highlights	24.02	16.62	19.74	15.61	3.17*	66	0.27
Negative Essay Highlights	17.27	17.55	18.98	16.09	-1.02	66	-0.10
Positive FWB	-2.68	2.20	-2.22	1.66	-1.70	50	-0.24
Negative FWB	.01	.79	.07	.64	-.54	50	-0.08
Constructive FWB	-2.16	2.70	-1.78	2.46	-.90	50	-0.15

Note: N = 68, Two-tailed critical value is 1.96, Bonferroni Correction for 8 comparisons is .01
Feedback Withholding Bias (FWB) – difference between written and spoken feedback
* $p < .01$

The Hayes (2009) Process Model program was utilized to test Hypothesis 3 that greater stereotype threat will be associated with greater levels of the FWB. This analysis program allows for multiple mediators to be included in one model, thus all four measures of stereotype threat (i.e., Stroop (working memory indicator), anxiety emotions, pulse rate and MAP) were included as parallel mediators to predict each of the different FWB measures (with each FWB measure in a different model), producing a total of 5 models. The IV was dichotomous such that 1 indicates when participants provided feedback to the White student and 2 indicates when participants provided feedback to the Black student. Essay and race order were again included as control variables and modeled as covariates.

The vast majority of path estimates in these models were not significant. The only model that significantly predicted one of the FWB indicators was for the outcome variable of Positive Essay Highlights ($F(5,128) = 2.87$, $p = .02$). In this model, student race was related to higher levels of MAP ($F(1,132) = 6.61$, $p = .01$),

and pulse rate ($F(5,128) = 2.87$, $p = .02$) had a positive predictive relationship with Positive Essay Highlights. Perhaps more importantly, student race had a marginally significant positive direct effect on Positive Essay Highlights such that Black students received more positive highlights than White students (see Figure 8 for estimates related to positive essay highlights and Table 4 for analysis data with positive highlights, see Tables 9 – 12 in Appendix C for results of remaining four nonsignificant models). Additional analyses with one mediator per model, rather than parallel mediators in one model, produced similar results to the parallel mediator models. As no significant indirect effects through the mediators were found, Hypothesis 3 was not supported.

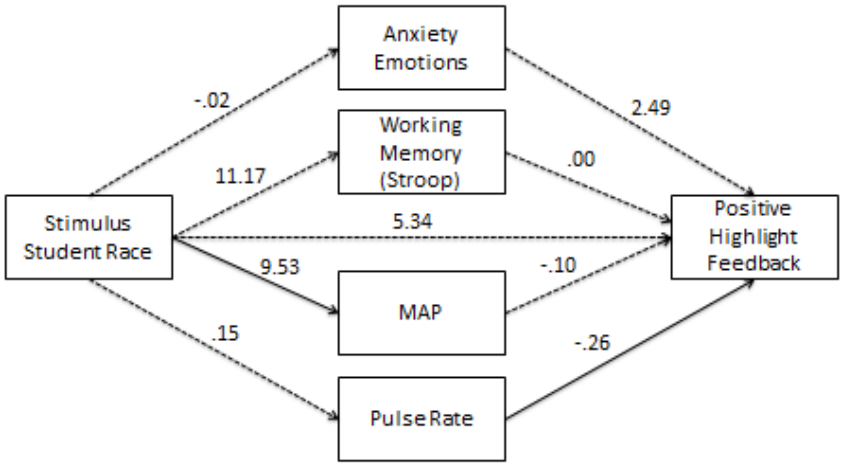

Figure 8. Hypothesis 3 analysis of stereotype threat indicators mediation effects on positive essay highlights. Unstandardized regression coefficients displayed.
Note. $p < .05$;Significant paths marked by solid line. Nonsignificant paths marked by dotted line.

Table 4

Mediation analysis with student race as IV, the four stereotype threat measures as parallel mediators, and positive essay highlights as the DV.

	Consequent (Difference in Positive Highlights)														
	M_1 (Anxiety)			M_2 (Stroop)			M_3 (MAP)			M_4 (Pulse)			Y (Positive Highlights)		
Antecedent	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*
X (FB Race)	-.02	.12	.89	11.17	64.23	.86	9.53	3.71	.01	.15	2.34	.95	5.34	2.78	.06
M_1 (Anxiety)													2.49	1.98	.21
M_2 (Stroop)													.00	.00	.97
M_3 (MAP)													-.10	.07	.12
M_4 (Pulse)													-.26	.10	.01
Constant	1.95	.19	.00	1066.03	101.53	.00	-8.39	5.87	.15	1.16	3.69	.75	9.87	6.72	.14
	R^2 = .00			R^2 = .00			R^2 = .05			R^2 = .00			R^2 = .10		
	F(1,132)=.02, *p*=.89			*F*(1,132)=.03, *p*=.86			*F*(1,132)=6.61, *p*=.01			*F*(1,132)=.00, *p*=.95			*F*(5,128)=2.87, *p*=.02		

The Hayes (2009) Process Model for moderation was used to test Hypothesis 4, which predicted Internal MCP would moderate the association between External MCP and the different indicators of the FWB when providing feedback to a Black student. One model was significant, with results indicating Internal MCP moderates the relationship between External MCP and the Negative FWB (see Figure 9 for graphical results). Participants low in Internal MCP yet high in External MCP were more likely to exhibit a negative FWB (in which they wrote more negative comments in their private notes than they conveyed to the student), however the effect becomes non-significant as levels of Internal MCP rise. Internal MCP failed to moderate the relationship between External MCP and the other FWB indicators including Positive Highlights, Negative Highlights, Positive FWB, and Constructive FWB (see Table 5). Thus, Hypothesis 4 was partially supported.

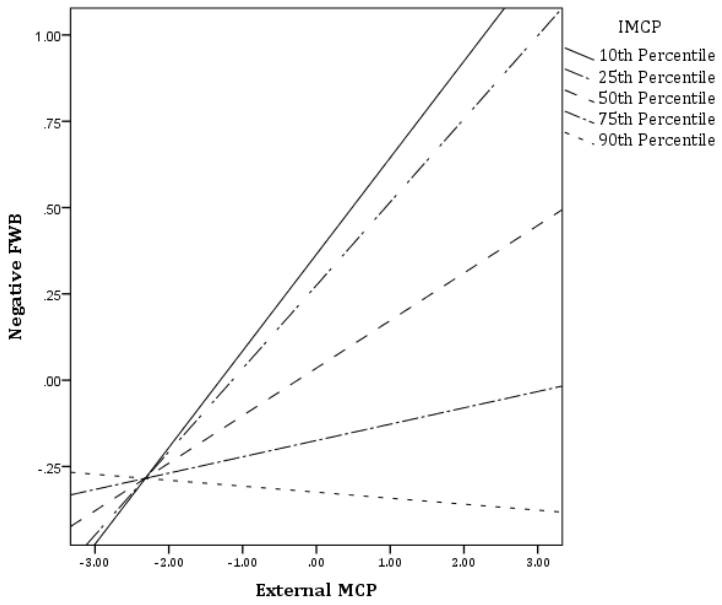

Figure 9. The moderation of Internal MCP on the relationship between External MCP and the Negative FWB indicator in the experimental (i.e., Black student) condition.

Table 5

Moderation of Internal MCP (IMCP) on the relationship between External MCP (EMCP) and FWB indicators in the experimental (i.e., Black student) condition

		Coeff.	*SE*	*t*	*p*
Model 1: Positive Highlights					
R^2 = .04, *MSE* = 279.39					
Intercept	i_1	24.11	2.10	11.49	.00
EMCP (*X*)	b_1	.16	1.24	.13	.90
IMCP (*M*)	b_2	1.65	1.08	1.52	.13
EMCP x IMCP	b_3	.10	.46	.21	.83
Model 2: Negative Highlights					
R^2 = .03, *MSE* = 318.52					
Intercept	i_1	17.11	2.24	7.64	.00
EMCP (*X*)	b_1	-1.60	1.32	-1.21	.23
IMCP (*M*)	b_2	.23	1.56	.20	.84
EMCP x IMCP	b_3	.04	.49	.08	.94
Model 3: Positive FWB					
R^2 = .03, *MSE* = 4.68					
Intercept	i_1	-2.71	.30	-8.94	.00
EMCP (*X*)	b_1	-.15	.19	-.78	.44
IMCP (*M*)	b_2	-.09	.17	-.56	.58
EMCP x IMCP	b_3	-.03	.09	-.36	.72
Model 4: Negative FWB					
R^2 = .31, *MSE* = .40					
Intercept	i_1	.01	.09	.15	.88
EMCP (*X*)	b_1	**.14**	**.05**	**2.58**	**.01**
IMCP (*M*)	b_2	**-.17**	**.05**	**-3.39**	**.00**
EMCP x IMCP	b_3	**-.06**	**.03**	**-2.52**	**.02**
Model 5: Constructive FWB					
R^2 = .01, *MSE* = 7.40					
Intercept	i_1	-2.10	.38	-5.51	.00
EMCP (*X*)	b_1	-.15	.23	-.63	.53
IMCP (*M*)	b_2	.12	.21	.56	.56
EMCP x IMCP	b_3	.02	.11	.18	.86

Model 1 – 2: *N* = 66; Model 3 – 5: *N* = 53

Multiple process models (Hayes, 2009) were utilized to test the moderated mediation predicted in Hypothesis 5, which posited that when providing feedback to the Black student Internal MCP would moderate the relationship between External MCP and the stereotype threat measures, which were entered as parallel mediators, predicting the FWB indicators. Results support the moderated mediation model predicting positive essay highlights. Model results show that the interaction between Internal MCP and External MCP explained a significant amount of variance in anxiety emotions reported, $R^2 = .26$, $F(5,60) = 4.27$, $p < .01$. See Table 6 for full results and Figure 10 for a graphical representation of the interaction with anxiety emotions. Results suggest that those participants reporting higher levels of both External MCP and Internal MCP report higher levels of anxiety emotions when providing feedback to Black students, however anxiety emotions failed to function as a mediator of positive highlights.

Results also support Internal MCP moderating the relationship External MCP and working memory (i.e., Stroop) in stereotype threat, $R^2 = .41$, $F(5,60) = 8.41$, $p < .01$, where participants with high External MCP and low Internal MCP display a more depleted working memory capacity. See Figure 11 for a graphical representation of the interaction.

The mediation analysis indicates that participant pulse rate has a significant positive predictive relationship with essay highlights for the Black student's essay such that participants with higher pulse rates made fewer positive highlights for Black students (see Figure 12 for model of results). No direct effect of External MCP on positive essay highlights was found, however there was evidence of an indirect

effect of External MCP on positive highlights at low levels of Internal MCP indicating a limited moderated mediation relationship (see Table 7 for conditional indirect effects). All other moderation mediation models were nonsignificant (see Tables 13 – 16 in Appendix C for full analysis data). Thus, Hypothesis 5 is only partially supported.

Table 6

Moderated mediation analysis with External MCP as the IV, Internal MCP as the moderator, four stereotype threat measures as parallel mediators, and positive essay highlights as the DV.

Antecedent		M_1 (Anxiety)			*M₂* (Stroop)			*M₃* (MAP)			*M₄* (Pulse Rate)			Y (Pos. Highlights)		
		Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*
X (EMCP)	a_1	-.08	.05	.08	16.97	20.78	.42	.28	1.79	.88	.44	.85	.61	c_1 .82	1.12	.47
M_1														b_1 2.17	3.00	.47
M_2														b_2 .00	.01	.99
M_3														b_3 -.11	.09	.20
M_4														***b₄*** **-.59**	**.17**	**.00**
V (IMCP)	a_2	.06	.04	.17	**-44.24**	**18.42**	**.02**	-.71	1.59	.66	-.89	.76	.24			
IMCP x EMCP	i_1	**-.05**	**.02**	**.00**	**-21.23**	**7.61**	**.01**	-.50	.66	.45	-.39	.31	.22			
Constant		**2.91**	**.35**	**.00**	**1838.62**	**151.72**	**.00**	**35.89**	**13.06**	**.01**	**12.89**	**6.23**	**.04**	22.61	15.80	.16
		R^2 =.26			R^2 = .41			R^2 = .07			R^2 =.09			R^2 =.22		
		F(5,60)=4.27, *p*=.00			*F*(5,60)=8.41, *p*=.00			*F*(5,60)=.92, *p*=.48			*F*(5,60)=1.22, *p*=.31			*F*(7,58)=2.29, *p*=.04		

N = 66; Control variables: Order of essay and race of student

Table 7

Conditional indirect effect(s) of EMCP on Positive Essay Highlights at different values of the moderator IMCP

IMCP Level	Anxiety Emotions		Stroop		MAP		Pulse Rate	
	Coeff.	95% Bias-Corrected Bootstrap CI	Coeff.	95% Bias-Corrected Bootstrap CI	Coeff.	95% Bias-Corrected Bootstrap CI	Coeff.	95% Bias-Corrected Bootstrap CI
-1.99	.04	-.16 –.64	-.00	-.99 – 1.07	-.14	-1.12 –.20	**-.71**	**-2.06 – -.06**
0.00	**-.18**	**-1.07 –.26**	-.00	-.44 –.48	-.03	-.73 –.52	-.26	-1.87 –.48
1.99	-.41	-1.78 –.74	.00	-.57 –.57	.08	-.56 – 1.41	.19	-1.71 – 1.78

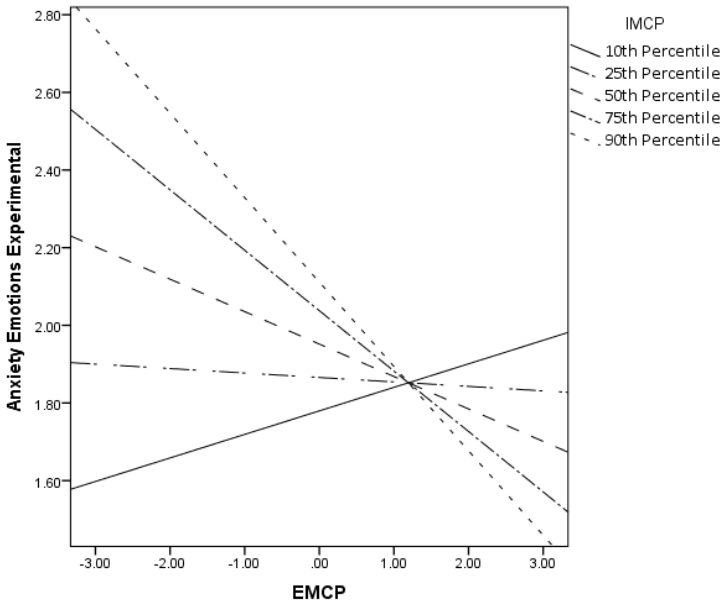

Figure 10. Interaction of participants' Internal MCP (EMCP) and Internal MCP (IMCP) with the stereotype threat indicator anxiety emotions.

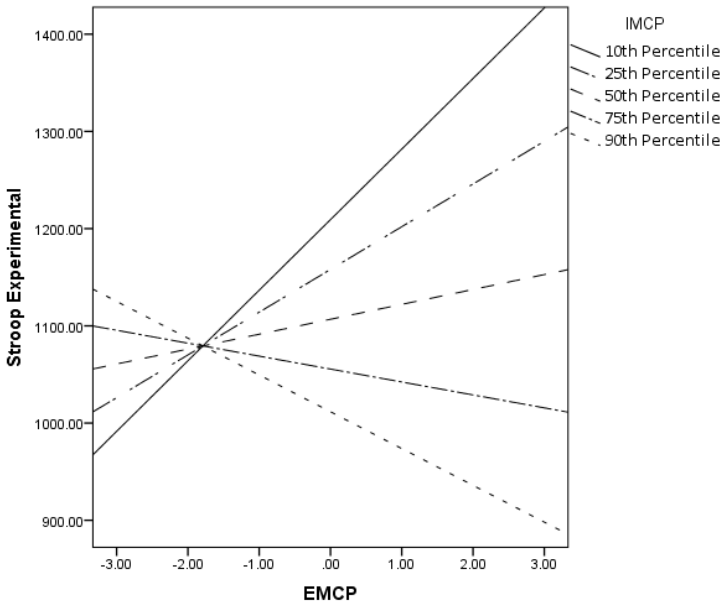

Figure 11. Interaction of participants' Internal MCP (EMCP) and Internal MCP (IMCP) with the working memory (Stoop) stereotype threat indicator.

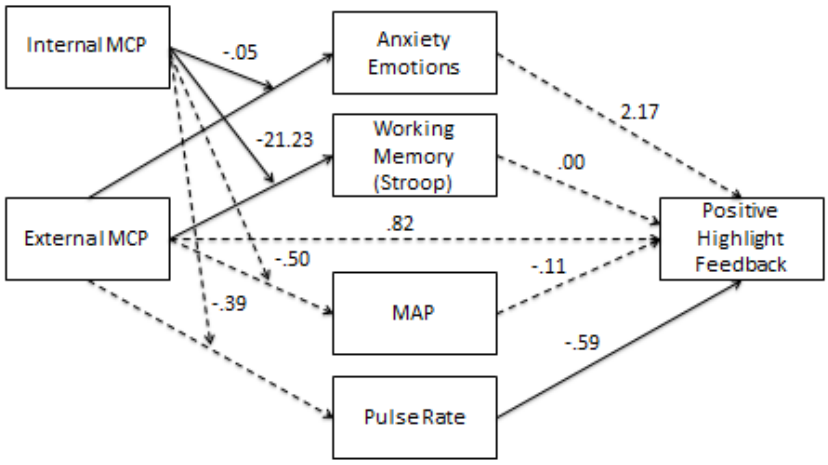

Figure 12. Model of moderated mediation analysis with External MCP as the IV, Internal MCP as the moderator, four stereotype threat measures as parallel mediators, and positive essay highlights as the DV. Unstandardized regression coefficients displayed.
Note. $p < .05$; Significant paths marked by solid line. Nonsignificant paths marked by dotted line.

Hypothesis 6 was tested using the Hayes Process Model for moderation to estimate the moderation of Implicit MCP (i.e., Belief One Is Prejudiced (BOP) and Negative Attitudes towards Prejudice (NAP) with the stereotype threat indicators. NAP was entered as the IV with BOP as the moderating variable between NAP and the different stereotype threat indicators. No evidence was found to support a direct or interactive relationship between NAP or BOP and any of the four stereotype threat indicators (see Table 8). Therefore, Hypothesis 6 was not supported.

Table 8

Moderation of Belief One's Prejudiced (BOP) on the relationship between Negative Attitudes toward Prejudice (NAP) and stereotype threat indicators

		Coeff.	*SE*	*t*	*p*
Model 1: Anxiety Emotions					
R^2 = .19, *MSE* = .47					
Intercept	i_1	2.90	.39	7.43	.00
NAP (*X*)	b_1	.37	.26	1.31	.16
BOP (*M*)	b_2	.34	.26	1.31	.20
NAP x BOP	b_3	.07	.59	.12	.90
Model 2: Stroop					
R^2 = .27, *MSE* =96677.36					
Intercept	i_1	1780.13	176.98	10.06	.00
NAP (*X*)	b_1	65.91	117.63	.56	.58
BOP (*M*)	b_2	59.53	117.49	.51	.61
NAP x BOP	b_3	-70.17	267.66	-.27	.79
Model 3: MAP					
R^2 = .07, *MSE* = 537.74					
Intercept	i_1	29.56	13.20	2.24	.03
NAP (*X*)	b_1	-11.05	8.77	-1.26	.21
BOP (*M*)	b_2	4.99	8.76	.57	.57
NAP x BOP	b_3	21.02	19.96	1.05	.30
Model 4: Pulse Rate					
R^2 = .05, *MSE* = 201.12					
Intercept	i_1	7.89	8.07	.98	.33
NAP (*X*)	b_1	6.60	5.36	1.23	.22
BOP (*M*)	b_2	-3.37	5.36	-.63	.53
NAP x BOP	b_3	8.17	12.21	.67	.51

N = 65

Five separate Hayes (2009) Process Models for moderated mediation were utilized to Hypothesis 7 that stereotype threat mediates the relationship between Implicit MCP (i.e., BOP and NAP) and the FWB indicators when providing feedback to the Black student. Of the five models, no evidence was found to support the hypothesis (see Tables 17 – 21 in Appendix C for results). That is, there is no

evidence to support Hypothesis 7 of a moderation relationship between Implicit MCP and the stereotype threat indicators as well as no evidence of the stereotype threat indicators mediating the relationship between Implicit MCP and the FWB indicators.

Additional Analyses

In order to fully explore the data, additional analyses were performed. Croft and Schmader (2012) failed to find differences between groups in letter grade assigned, and the current study's results replicate this result. A paired samples *t*-test with a Bonferroni correction was conducted to assess the difference in ratings on letter grade and an original outcome variable of whether the student should receive the scholarship. The group differences did not reach significance for either outcome variable. Results do suggest that despite participants assigning similar letter grades to both the Black and White student (t (68) = 1.41, p = .16), they overall rated the Black student (M = 3.19) as more qualified than the White student (M = 2.91) to receive the scholarship (t (68) = 2.27, p = .027). Due to the Bonferroni correction for two comparisons, this difference fails to reach the level of statistically significant.

Discussion

The importance of constructive feedback to learning and performance has been established in previous research (see Barron, 1988; Jussim & Eccles, 1992). The results of this research do not support the proposal that Black students are likely to receive less constructive feedback than White students. Participants provided similar amounts of constructive feedback when they believed the feedback

recipient was Black as when they believed the feedback recipient was White. The results do suggest that Black students received more positive feedback compared to White students from these White evaluators. In other words, participants seemed to "cushion" their feedback with more positive comments.

Although there was some support for stereotype threat occurring concurrently with providing feedback, evidence does not support that stereotype threat functions as a mechanism for feedback bias. Specifically, this study supports that White evaluators experience greater stereotype threat emotions (i.e., anxiety, tension) when providing feedback to a Black student. In addition, a predictive relationship with the stereotype threat emotions and working memory depletion and explicitly measured Motivation to Control Prejudice (i.e., Internal MCP and external MCP). Support was not found for a physiological stereotype threat reaction, which will be discussed with more detail in the Limitations section. The measures of stereotype threat failed to function as mediators between student race and feedback bias and therefore stereotype threat is not supported as the mechanism by which interracial feedback bias occurs.

Some support was found for a negative FWB where External MCP and Internal MCP interacted to predict participants writing more negative feedback comments in the "private" notes than spoken during the recorded feedback. In line with previous research (Croft & Schmader, 2012), individuals motivated to *appear* nonprejudiced without a concurrent motivation to actually *be* nonprejudiced were more likely to withhold negative feedback, (but not positive or constructive feedback). Negative feedback was operationalized as critical feedback without

suggestion for improvement. It is possible that participants provided suggestions for improvement when giving recorded feedback that they did not include in the notes section and it is also possible that they simply omitted the negative written comments in the recorded feedback. Participants were less likely to show a negative FWB at higher levels of Internal MCP suggesting that the motivation to appear nonprejudiced was a driving factor in displaying the negative FWB.

The results of the current study also provide evidence of an interracial feedback bias where White evaluators provide more positive feedback to Black students compared to White students. This finding replicates previous studies (Biernat & Danahar, 2012; Harber, 1998; Harber et al., 2010) that found that White participants had a positive feedback bias when providing feedback to a racial minority. In past work, a potential explanation for the provision of overly positive feedback was that the work of minority students was compared against a lower standard than the work of white students (e.g., Biernat & Manis, 1994). Thus, the stereotypes suggesting that people of color have lower academic ability may result in lower expectations for student performance. In such cases, mediocre work of minority students may be interpreted more positively than if the same quality work had been performed by White students (see Harber, 1998). Several studies have suggested that these types of processes, as described by the Shifting Standards Model, operate in a variety of domains (see Biernat, Manis, & Nelson, 1991; Biernat & Danahar, 2012; Harber, 1998).

A small body of research, based on the Shifting Standards Model (Biernat et al., 1991), has consistently found that subjective feedback can, in fact, be impacted

by stereotypes, such that common stereotypes are used as a standard for comparison or evaluation (see Biernat et al., 1991). That is, when arriving at subjective evaluations, there is a general tendency to compare the target with others from the same racial group rather than to evaluate successive targets against a common, unchanging set of standards (Biernat, 2012; Biernat & Danahar, 2012). It may be possible that feedback providers are shifting their standards for Black students (Biernat & Danahar, 2012) and perhaps feedback providers are making the assumption the Black students do not expect or will not utilize constructive feedback. Similar to past research, participants in this study recommended both the Black and White students at similar levels and provided similar grades across the essays (Kobrynowicz & Biernat, 1997). Those results coupled with the more positive subjective feedback for Black students fits the Shifting Standards model where participants are more likely to display bias on objective measures but not display bias, or show a pro-minority bias on subjective measures.

Thus, the lack of constructive feedback from White evaluators would not be related to stereotype threat, but differing feedback expectations or standards for Black students compared to White students. In fact, for many years, researchers observing the phenomenon of overly positive feedback were not able to determine whether it was caused by stereotype threat or shifting standards (Harber, 1998; Harber et al., 2010). The results of the current study fail to fully support the stereotype threat theory, even though there is evidence of stereotype threat occurring within the interracial feedback situation but not as a causal mechanism of the FWB.

Although the results of the current study support that Black students receive similar amounts of crucial, constructive criticism from White evaluators, the positive feedback bias present may have unintended implications for the feedback receiver. An early research study found that Black individuals discounted positive, subjective interpersonal feedback from White evaluators only when they believed the White evaluator knew they were Black (Crocker et al., 1991). This reaction is based on the commonly held belief that White people will attempt to avoid appearing prejudiced (Crocker et al., 1991; Dovidio, Kawakami, & Gaertner, 2002). Subsequent research indicates that when Black students received positive academic performance feedback (i.e., received a score and the phrase "Great Job!!"), they believed the White evaluator possessed lower expectations compared to Black students that did not receive praise (i.e., received a score only) on their academic performance (Lawrence et al., 2011). Thus, the impact of the positive feedback bias may be detrimental to the learning of minority students, as would be the FWB.

Croft and Schmader (2012) argue that stereotype threat is the underlying mechanism of the FWB based on results that show impression management concerns were the greatest predictor of the FWB. Despite the lack of evidence for stereotype threat as the FWB mechanism, the results still support the role of impression management in interracial feedback. External and Internal MCP were predictive of negative FWB, stereotype threat emotions, and depletion of working memory. Particularly important was the interactive role of External MCP, which measures motivation to not be perceived as prejudiced by others, with Internal MCP to predict measures of stereotype threat and the FWB. The relationship with MCP

and these outcome variables support that impression management has a role in experience of providing interracial feedback for White evaluators. Thus, the present study replicates Croft and Schmader's finding on impression management, but it fails to support the argument that stereotype threat is the mechanism of feedback bias.

Potential Limitations

This study has several potential limitations that should be noted. First, the study was conducted in a laboratory in which video recording and physiological assessment equipment were used. The use of video recording devices allowed the study to be conducted without confederates; however, this did reduce the ecological validity because the participants did not provide feedback with the recipient face-to-face. In addition, the laboratory component of the study may have reduced the recruitment of participants leading to a smaller sample size. The smaller sample size may have impacted the analyses power to detect significant relationships, especially in the more complex process models (Hayes, 2012).

Another possible issue may have been that the undergraduate participants may lack the knowledge to provide constructive feedback on the stimulus materials. Although the typical ACT score for students at the university is 26 – 31 (31 represents the 75th percentile for the ACT exam; US Department of Education, 2017) and participation requirements of being a STEM major and completing at least one semester of college courses (i.e., sophomore standing and above) were in place, it is possible that those requirements do not ensure participants had the ability to provide constructive feedback. Additionally, the perception among students that the

Missouri S&T campus is more selective and better rated for the STEM fields than UMSL may have influenced the participants' perceptions of the UMSL students. UMSLs' reputation with Missouri S&T students may have led to the participants to have lower expectations for the UMSL students. It may also limit the generalizability of these results as the relationship between state campuses is likely unique and may have uniquely affected the results of this study.

Another study limitation may be the region of the United States where data was collected. The St. Louis area and University of Missouri system has received increased media attention on race relations in recent years following the death of Michael Brown in August of 2015 and protests at the University of Missouri - Columbia. The focus of national media attention on negative race relations may have increased the salience of race in this study and possibly skewed results. Thus, it is possible that the increased attention to interracial tensions may have impacted this study in uncontrollable ways. Although there is no evidence of an impact, unfortunately there is also no way to investigate the possible impact.

This study did not use blatant stereotype threat cues (e.g., telling women about the stereotype women are bad drivers and then asking them to perform a driving test), but relied on subtle cues in the context to induce stereotype threat (see Murphy et al., 2007; Stone & McWhinnie, 2008). The methods used in past studies were more overt (e.g., Osborne, 2007). Thus, the lack of evidence may be related to individuals not perceiving the feedback as a stereotype threat situation which impacted the results of the study.

Finally, the physiological measurement tools utilized did not allow for continuous, precise measurements of blood pressure and pulse rate. The monitor utilized automatically reported blood pressure every 5 minutes, although the monitor was manually activated to take a measurement before the video recorded feedback, the inability to record continuously during the feedback likely hindered the ability to accurately assess the physiological stereotype threat component. Moreover, the equipment only allowed for the collection of blood pressure and pulse rate, of which both increase during positive and negative stress events. Therefore, this would not allow for an accurate interpretation of if the physiological arousal would be related to a challenge reaction which is adaptive (i.e., SAM axis) or to a threat response (e.g., stereotype threat) that is maladaptive (i.e., SAM and HPA axis).

Future Research

One area for future research would be to replicate the study using more advanced and precise physiological measurements. This would allow for better evidence to assess if all of the stereotype threat measurements are present in cross racial feedback situations, even if the relationship with FWB remains null. The investigation into stereotype threat's relationship with other feedback related outcomes, such as feedback acceptance, could provide additional information on ensuring effective interracial feedback.

Another area of future research is to investigate the relationship of past interracial interactions with participants' anxiety around providing interracial feedback. It may be that the participants in this study, especially given the vast

majority of students on the Missouri University of Science and Technology campus are White, have had limited interactions with Black students thus leading to more anxiety surrounding the feedback situation. It may be possible that the anxiety differences found in this study were also related to limited interracial interactions as well as impression management.

The order effects found with the stereotype threat measurements warrant future investigation. Individuals that provided feedback to the Black student first were more likely to demonstrate behavior that was consistent with the hypothesized relationships than participants that provided feedback to the Black student second. Those individuals who provided feedback to the White student first were more likely to have more consistent levels of the stereotype threat measures across both feedback sessions. Given these results, the relationship between feedback order and self-regulation for White evaluators may be an area to explore. Past research has found that individuals tend to demonstrate self-regulation in interracial interactions (Richeson & Shelton, 2007) and participants in this study may have recognized the interracial component of the study and engaged in self-regulation during the second feedback session with the Black student. It is also possible that the order effects represent a demand characteristic where the participants recognized the racial comparison and therefore demonstrated a response bias in the second feedback session with the Black student.

Conclusion

In general, the results of this study suggest that Black students do receive similar amounts of constructive feedback from White evaluators and are likely to

receive slightly more positive feedback than White students. Although impression management appears to impact slightly, this study suggests that Black students, specifically in the STEM fields, are likely to receive feedback that will allow them to learn and develop. The use of students in this study may also suggest that interracial peer feedback is likely to be similar to same race feedback for this generation of students. Therefore, this study presents positive implications for racial inclusion in learning within the STEM disciplines.

References

Alter, A. L., Aronson, J., Darley, J. M., Rodriguez, C., & Ruble, D. N. (2010). Rising to the threat: Reducing stereotype threat by reframing the threat as a challenge. *Journal of Experimental Social Psychology, 46*(1), 166-171. doi:10.1016/j.jesp.2009.09.014

Amodio, D. M., Harmon-Jones, E., Devine, P. G., Curtin, J. J., Hartely, S. L., & Covert, A. E. (2004). Neural signals for the detection of unintentional race bias. *Psychological Science, 15,* 88-93. doi: 10.1111/j.0963-7214.2004.01502003.x

Amodio, D. M., Kubota, J. T., Harmon-Jones, E., & Devine, P. G. (2006). Alternative mechanisms for regulating racial responses according to internal vs. external cues. *Social Cognitive and Affective Neuroscience, 1,* 26-36. doi:10.1093/scan/nsl002

Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When White men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35,* 29-46.

Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken, Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361-382). New York, NY, US: Guilford Press.

Baron, R. A. (1988). Negative effects of destructive criticism: Impact on conflict, self-efficacy, and task performance. *The Journal of Applied Psychology, 73*(2), 199-207. doi:10.1037/0021-9010.73.2.199

Baron, R. A. (1990). Countering the effects of destructive criticism: The relative efficacy of four interventions. *Journal of Applied Psychology, 75*(3), 235-245. doi:10.1037/0021-9010.75.3.235

Beede, D. N., Julian, T. A., Khan, B., Lehrman, R., McKittrick, G. Langdon, D., & Doms, M. E., (September 28, 2011). *Education Supports Racial and Ethnic Equality in STEM*. Economics and Statistics Administration Issue Brief No. 05-11. Retrieved from http://ssrn.com/abstract=1934821 or http://dx.doi.org/10.2139/ssrn.1934821

Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: Mechanism, alleviation, and spillover. *Journal of Experimental Psychology: General, 136,* 256-276. doi: 10.1037/0096-3445.136.2.256

Bell-Jordan, K. E. (2008). Black, White and a Survivor of the Real World: Constructions of race on reality TV. *Critical Studies in Media Communication, 25*, 353–372.

Bergsieker, H. B., Shelton, J. N., & Richeson, J. A. (2010). To be liked versus respected: Divergent goals in interracial interactions. *Journal of Personality and Social Psychology, 99*(2), 248-264. doi:10.1037/a0018474

Biernat, M., & Danahar, K. (2012). Interpreting and reacting to feedback in stereotype-relevant performance domains. *Journal of Experimental Social Psychology, 48,* 271-276. doi: 10.1016/j.jesp.2011.08.010

Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology, 66*(1), 5-20. doi:10.1037/0022-3514.66.1.5

Bjorklund, S. A., Parente, M., & Sathianathan, D. (2004). Effects of faculty interaction and feedback on gains in student skills. *Journal of Engineering Education, 92,* 153-160.

Blascovich, J., Spencer, S. J., Quinn, D., & Steele, C. (2001). African americans and high blood pressure: The role of stereotype threat. *Psychological Science, 12*(3), 225-229. doi:10.1111/1467-9280.00340

Blau, P. M. (1964). *Exchange and power in social life*. New York: J. Wiley.

Bogle, D. (2002). *Toms, coons, mulattoes, mammies, and bucks: An interpretive history of blacks in American films*. New York: Continuum.

Bosson, J. K., Haymovitz, E. L., & Pinel, E. C. (2004). When saying and doing diverge: The effects of stereotype threat on self-reported versus non-verbal anxiety. *Journal of Experimental Social Psychology, 40*(2), 247-255. doi:10.1016/S0022-1031(03)00099-4

Britt, T. W., & Crandall, C. S. (2000). Acceptance of feedback by the stigmatized and non-stigmatized: The mediating role of the motive of the evaluator. *Group Processes & Intergroup Relations*, *3*(1), 79-95.

Butz, D. A., & Plant, E. A. (2009). Prejudice control and interracial relations: The role of motivation to respond without prejudice. *Journal of Personality, 77*(5), 1311. doi:10.1111/j.1467-6494.2009.00583.x

Chang, S. H., & Kleiner, B. H. (2003). Common racial stereotypes. *Equal Opportunities International, 22*(3), 1-9. doi:10.1108/02610150310787388

Chartrand, T. L., & Bargh, J. A. (1996). Automatic activation of impression formation and memorization goals: Nonconscious goal priming reproduces effects of explicit task instructions. *Journal Of Personality And Social Psychology*, *71*(3), 464-478. doi:10.1037/0022-3514.71.3.464

Chen, X., & Thomas, W. (2009). *Students who study science, technology, engineering and mathematics (STEM) in post-secondary education.* NCES 2009-161. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

Chung, B. G., Ehrhart, M. G., Ehrhart, K., Hattrup, K., & Solamon, J. (2010). Stereotype threat, state anxiety, and specific self-efficacy as predictors of promotion exam performance. *Group & Organization Management, 35*(1), 77-107. doi:10.1177/1059601109354839

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, N.J: L. Erlbaum Associates.

Cohen, G. L., & Steele, C. M. (2002). A barrier of mistrust-chapter 15:How negative stereotypes affect cross-race mentoring. (pp. 303-327) Academic Press. doi:10.1016/B978-012064455-1/50018-X

Cohen, G. L., Steele, C. M., & Ross, L. D. (1999). The mentor's dilemma: Providing critical feedback across the racial divide. *Personality and Social Psychology Bulletin, 25,* 1302-1316. doi: 10.1177/0146167299258011

Cole, D. (2008). Constructive criticism: The role of student-faculty interactions on African American and Hispanic students' educational gains. *Journal of College Student Development, 49,* 587-605

Croft, A., & Schmader, T. (2012). The feedback withholding bias: Minority students do not receive critical feedback from evaluators concerned about appearing racist. *Journal of Experimental Social Psychology, 48,* 1139-1144. doi: 10.1016/j.jesp.2012.04.010

Croizet, J. C., & Claire, T. (1998). Extending the concept of stereotype and threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin, 24,* 588-594. doi: 10.1177/0146167298246003

Croizet, J., Després, G., Gauzins, M., Huguet, P., Leyens, J., & Méot, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality & Social Psychology Bulletin, 30*(6), 721-731. doi:10.1177/0146167204263961

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter? *The American Economic Review, 95*(2), 158-165. doi:10.1257/000282805774670446

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*(1), 5-18. doi:10.1037/0022-3514.56.1.5

Diekman, A. B., & Hirnisey, L. (2007). The effect of context on the silver ceiling: A role congruity perspective on prejudiced responses. *Personality and Social Psychology Bulletin, 33,* 1353-1366. doi: 10.1177/0146167207303019

Dipboye, R. L. (1985). Some neglected variable in research on discrimination in appraisals. *Academy of Management Review, 10,* 116-127.

Dovidio, J. F., & Gaertner, S. L. (2000). Aversive racism and selection decisions: 1989 and 1999. *Psychological Science, 11,* 319 – 323.

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2000). Reducing contemporary prejudice: Combating explicit and implicit bias at the individual and intergroup level. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp.137-164). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Dunton, B. C., & Fazio, R. H. (1997). An individual difference measure of motivation to control prejudiced reactions. *Personality and Social Psychology Bulletin, 23,* 316–326.

Eagly, A. H., & Carli, L. L. (2007). *Through the labyrinth: The truth about how women become leaders.* Boston: Harvard Business School Press.

Eagly, A. H., & Chin, J. L. (2010). Diversity and leadership in a changing world. *American Psychologist, 65,* 216-224. doi: 10.1037/a0018957

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109,* 573-598. doi: 10.1037/0033-295x.109.3.573

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19-23. doi:10.1111/1467-8721.00160

Eysneck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion, 6,* 409-434.

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 357-411). New York: McGraw-Hill.

Frantz, C. M., Cuddy, A. J. C., Burnett, M., Ray, H., & Hart, A. (2004). A threat in the computer: The race implicit association test as a stereotype threat experience. *Personality and Social Psychology Bulletin, 30,* 1611-1624. doi: 10.1177/0146167204266650

Glaser, J., & Knowles, E. D. (2008). Implicit motivation to control prejudice. *Journal of Experimental Social Psychology, 44*(1), 164-172. doi:10.1016/j.jesp.2007.01.002

Goff, P., Steele, C. M., & Davies, P. G. (2008). The space between us: Stereotype threat and distance in interracial contexts. *Journal of Personality And Social Psychology*, *94*(1), 91-107. doi:10.1037/0022-3514.94.1.91

Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of Latino women. *Personality and Social Psychology Bulletin, 28,* 659-670. doi: 10.1177/0146167202288010

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the implicit association test: III. meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*(1), 17-41. doi:10.1037/a0015575

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology, 39*(3), 281-291. doi:10.1017/S0048577201393198

Harber, K. D. (1998). Feedback to minorities: Evidence of a positive bias. *Journal of Personality and Social Psychology, 74,* 623-628.

Harber, K. D. (2004). The positive feedback bias as response to out-group unfriendliness. *Journal of Applied Social Psychology, 34,* 2272-2297.

Harber, K. D., Gorman, J. L., Gengaro, F. P., Butisingh, S., Tsang, W., & Ouellette, R. (2012). Students' race and teachers' social support affect the positive feedback bias in public schools. *Journal Of Educational Psychology*, *104*(4), 1149-1161. doi:10.1037/a0028110

Harber, K. D., Stafford, R., & Kennedy, K. A. (2010). The positive feedback bias as a response to self-image threat. *The British Journal of Social Psychology, 49,* 207-218.

Hayes, A. (2009). Beyond baron and kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76*(4), 408-420. doi:10.1080/03637750903310360

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Heilman, M. E. (1995). Sex stereotypes and their effects in the workplace: What we know and what we don't know. *Journal of Social Behavior and Personality, 10*, 3-26.

Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues, 57*(4), 657-674. doi:10.1111/0022-4537.00234

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52*(12), 1280-1300. doi:10.1037/0003-066X.52.12.1280

Higgins, E. T. (1999). Promotion and prevention as a motivational duality: Implications for evaluative processes. In S.Chaiken & Y.Trope (Eds.), *Dual-process theories in social psychology* (pp. 503– 525). New York: Guilford Press.

Ho, H. T. & Yeung, C. (2014). Giving feedback to clients. *Management Science, 60*(8), 1926-1944. doi:10.1287/mnsc.2013.1846

Jeffries, C. H., & Hornsey, M. J. (2012). Withholding negative feedback: Is it about protecting the self or protecting others?. *British Journal Of Social Psychology*,*51*(4), 772-780. doi:10.1111/j.2044-8309.2012.02098.x

Jones, B. D., Ruff, C., & Paretti, M. C. (2013). The impact of engineering identification and stereotypes on undergraduate women's achievement and persistence in engineering. *Social Psychology of Education, 16,* 471-493. doi: 10.1007.s11218-013-9222-x

Jussim, L., & Eccles, J. (1992). Teacher expectations II: Construction and reflection of student achievement. *Journal of Personality and Social Psychology, 63*(3), 947-961.

Klopfenstein, K. (2005). Beyond test scores: The impact of black teacher role models on rigorous math taking. *Contemporary Economic Policy, 23*(3), 416-428. doi:10.1093/cep/byi031

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*, 254-284.

Kobrynowicz, D., & Biernat, M. (1997). Decoding subjective evaluations: How stereotypes provide shifting standards. *Journal of Experimental Social Psychology*, *33*(6), 579-601.

Landivar, L. C. (September 2013). Disparities in STEM employment by sex, race, and Hispanic origin: American community survey reports. Retrieved from http://www.census.gov/prod/2013pubs/acs-24.pdf

Long, M., Boiarsky, G., & Thayer, G. (2001). Gender and racial counter-stereotypes in science education television: A content analysis. *Public Understanding of Science, 10*(3), 255-269. doi:10.1088/0963-6625/10/3/301

Mallinckrodt, B., Miles, J. R., Bhaskar, T., Chery, N., Choi, G., & Sung, M. R. (2014). Developing a comprehensive scale to assess college multicultural programming. *Journal of counseling psychology*, *61*(1), 133.

Martocchio, J. J., & Dulebohn, J. (1994). Performance feedback effects in training: The role of perceived controllability. *Personnel Psychology, 47*(2), 357-373. doi:10.1111/j.1744-6570.1994.tb01729.x

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*(3), 207-230. doi:10.1207/S15327043HUP1603_3

Mendes, W. B., Blascovich, J., Hunter, S. B., Lickel, B., & Jost, J. T. (2007). Threatened by the unexpected: Physiological responses during social interactions with expectancy-violating partners. *Journal of Personality and Social Psychology, 92*(4), 698-716. doi:10.1037/0022-3514.92.4.698

Mendes, W. B., Blascovich, J., Lickel, B., & Hunter, S. (2002). Challenge and threat during social interaction with White and Black men. *Personality and Social Psychology Bulletin, 28,* 939-952. doi: 10.1177/014616720202800707

Metcalfe, J., & Jacobs, W. J. (1998). Emotional memory: The effects of stress on “cool” and “hot” memory systems. In D. L. Medin (Ed.), *The psychology of learning and motivation: Vol 38. Advances in research and theory* (pp. 187-222). San Diego, CA: Academic Press.

Meyers, L. S., Gamst, G., & Guarino, A. J. (2006). *Applied Multivariate Research.* Thousands Oaks, CA; Sage Publications, Inc.

Monteith, M. J., Spicer, C. V., & Tooman, G. D. (1998). Consequences of stereotype suppression: Stereotypes on AND not on the rebound. *Journal of Experimental Social Psychology, 34*(4), 355-377. doi:10.1006/jesp.1998.1355

Moskowitz, G. B., Gollwitzer, P. M., Wasel, W., & Schaal, B. (1999). Preconscious control of stereotype activation through chronic egalitarian goals. *Journal of Personality and Social Psychology, 77*(1), 167-184. doi:10.1037/0022-3514.77.1.167

Moskowitz, G. B., & Li, P. (2011). Egalitarian goals trigger stereotype inhibition: A proactive form of stereotype control. *Journal of Experimental Social Psychology*, *47*(1), 103-116. doi:10.1016/j.jesp.2010.08.014

Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology, 89*(6), 852-863. doi:10.1037/0022-3514.89.6.852

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin, 126*(2), 247-259. doi:10.1037/0033-2909.126.2.247

Murphy, M., Steele, C., & Gross, J. (2007). Signaling threat: How situational cues affect women in math, science, and engineering. *Psychological Science, 18,* 879-885. doi: 10.1111/j.1467-9280.2007.01995.x

National Science Foundation (2011). *Women, Minorities, and Persons with Disabilities in Science and Engineering* (NSF Publication No. 13-304). Retrieved from http://www.nsf.gov/statistics/wmpd/2013/

New York City Open Data (2010). *Popular baby names by race/ethnicity (*Data file). Retrieved from https://data.cityofnewyork.us/City-Government/Most-Popular-Baby-Names-By-Race-Ethnicity/iyrc-jsee

Osborne, J. W. (2001). Testing stereotype threat: Does anxiety explain race and sex differences in achievement? *Contemporary Educational Psychology, 26*(3), 291-310. doi:10.1006/ceps.2000.1052

Park, S. H., Glaser, J., & Knowles, E. D. (2008). Implicit motivation to control prejudice moderates the effect of cognitive depletion on unintended discrimination. *Social Cognition, 26*(4), 401-419. doi:10.1521/soco.2008.26.4.401

Plant, E. A. (2004). Responses to interracial interactions over time. *Personality and Social Psychology Bulletin, 30*, 1458–1471.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal Of Personality And Social Psychology*, *75*(3), 811-832. doi:10.1037/0022-3514.75.3.811

Plant, E. A., & Devine, P. G. (2009). The active control of prejudice: Unpacking the intentions guiding control efforts. *Journal of Personality and Social Psychology, 96*(3), 640-652. doi:10.1037/a0012960

Plant, E. A., Devine, P.G., & Peruche, B.M. (2010). Regulatory Concerns for Interracial Interactions: Approaching Egalitarianism versus Avoiding Prejudice. *Personality and Social Psychology Bulletin, 36,* 1135-1147.

Potts, R., & Martinez, I. (1994). Television viewing and children's beliefs about scientists. *Journal of Applied Developmental Psychology*, *15*(2), 287-300.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods,*

Instruments, & Computers : A Journal of the Psychonomic Society, Inc, 36(4), 717-731. doi:10.3758/BF03206553

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. (2007). *Multivariate Behavioral Research, 42*(1), 185-227. doi:10.1080/00273170701341316

Price, M., Handley, K., Miller, J., & O'Donovan, B. (2010). Feedback: All that effort, but what is the effect? *Assessment and Evaluation in Higher Education, 35,* 277-289. doi: 10.1080/02602930903541007

Reyna, C. (2000). Lazy, dumb, or industrious: When stereotypes convey attribution information in the classroom. *Educational Psychology Review, 12*(1), 85-110. doi:10.1023/A:1009037101170

Richeson, J. A., & Shelton, J. N. (2007). Negotiating interracial interactions: Costs, consequences, and possibilities. *Current Directions in Psychological Science*, *16*(6), 316-320.

Robst, J. Keil, J., & Russo, D. (1998). The effect of gender composition of faculty on student retention. *Economics of Education Review, 17*(4), 429-439.

Rosenthal, H. E. S., & Crisp, R. J. (2006). Reducing stereotype threat by blurring intergroup boundaries. *Personality and Social Psychology Bulletin, 32*(4), 501-511. doi:10.1177/0146167205281009

Ruscher, J. B., Wallace, D. L., Walker, K. M., & Bell, L. H. (2010). Constructive feedback in cross-race interactions. *Group Processes and Intergroup Relations, 13,* 603-619.

Sawyer, P. J., Major, B., Casad, B. J., Townsend, S. S. M., & Mendes, W. B. (2012). Discrimination and the stress response: Psychological and physiological consequences of anticipating prejudice in interethnic interactions. *American Journal of Public Health, 102*(5), 1020-1026. doi:10.2105/AJPH.2011.300620

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated model of stereotype threat effects on performance. *Psychological Review, 115,* 336-356. doi: 10.1037/0033-295X.115.2.336

Schmader, T., Johns, M., & Barquissau, M. (2004). The costs of accepting gender differences: The role of stereotype endorsement in women's experience in the math domain. *Sex Roles, 50,* 835-850. doi: 10.1023/B:SERS.0000029101.74557.a0.

Schommer, N. C., Hellhammer, D. H., & Kirschbaum, C. (2003). Dissociation between reactivity of the hypothalamus-pituitary-adrenal axis and the sympathetic-adrenal-medullary system to repeated psychosocial stress. *Psychosomatic Medicine, 65,* 450-460.

Sczesny, S., & Stahlberg, D. (2002). The influence of gender-stereotyped perfumes on leadership attribution. *European Journal of Social Psychology, 32*(6), 815-828. doi:10.1002/ejsp.123

Seibt, B., & Forster, J. (2004). Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory foci. *Journal of Personality and Social Psychology, 87,* 38-56. doi: 10.1037/0022-3514.87.1.38

Sellers, R. M., & Shelton, J. N. (2003). The role of racial identity in perceived racial discrimination. *Journal of Personality and Social Psychology, 84*(5), 1079-1092. doi:10.1037/0022-3514.84.5.1079

Sherwood, A., Allen, M. T., Fahrenberg, J., Kelsey, R. M., Lovallo, W. R., & van Doornen, L. J. (1990). Methodological guidelines for impedance cardiography. *Psychophysiology, 27*(1), 1-23.

Spencer, S., Steele, C. M., & Quinn, D. (1999). Under suspicion of inability: Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35,* 4-28.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613-629.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of african americans. *Journal of Personality and Social Psychology, 69*(5), 797-811. doi:10.1037/0022-3514.69.5.797

Stone, J. (2002). Battling doubt by avoiding practice: the effects of stereotype threat on self-handicapping in White athletes. *Personality and Social Psychology Bulletin, 28,* 1667-1678. doi: 10.1177/014616702237648

Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on Black and White athletic performance. *Journal of Personality and Social Psychology, 77,* 1213-1227.

Stone, J., & McWhinnie, C. (2008). Evidence that blatant versus subtle stereotype threat cues impact performance through dual processes. *Journal of*

Experimental Social Psychology, 44(2), 445-452. doi:10.1016/j.jesp.2007.02.006

Tyree, T. (2011). African american stereotypes in reality television. *Howard Journal of Communications, 22*(4), 394-413. doi:10.1080/10646175.2011.617217

United States Department of Education (2017). CollegeScorecard. Retrieved from https://collegescorecard.ed.gov/school/?178411-Missouri-University-of-Science-and-Technology

von Hippel, W., von Hippel, C., Conway, L., Preacher, K. J., Schooler, J. W., & Radvansky, G. A. (2005). Coping With Stereotype Threat: Denial as an Impression Management Strategy. *Journal Of Personality And Social Psychology*, *89*(1), 22-35. doi:10.1037/0022-3514.89.1.22

Walls, L. (2012). Third grade African American students' views of the nature of science. *Journal of Research in Science Teaching, 49*(1), 1-37. doi:10.1002/tea.20450

Wang, Y. W., Davidson, M. M., Yakushko, O. F., Savoy, H. B., Tan, J. A., & Bleier, J. K. (2003). The scale of ethnocultural empathy: development, validation, and reliability. *Journal of counseling psychology*, *50*(2), 221.

Wegner, D. M., Erber, R., & Zanakos, S. (1993). Ironic processes in the mental control of mood and mood-related thought. *Journal of Personality and Social Psychology, 65*(6), 1093-1104. doi:10.1037/0022-3514.65.6.1093

Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behavior: A review of possible mechanisms. *Psychological Bulletin, 127*(6), 797-826. doi:10.1037/0033-2909.127.6.797

Yeung, N. C. J., & von Hippel, C. (2008). Stereotype threat increases the likelihood that female drivers in a simulator run over jaywalkers. *Accident Analysis and Prevention, 40,* 667-674. doi: 10.1016/j.aap.2007.09.003

Appendix A

Coding of Feedback Comments Information

Two coders, blind to the experimental conditions, were provided with all of the essay comments from the participants. Each coder used a standard rating form to code the positive content comments (e.g., you made a good point here), the negative content comments (e.g., how does this information relate?), constructive (e.g., work on transition out of introduction paragraph), destructive (e.g., you should work on your writing), and neutral comments (e.g., good luck)

The two coders practiced on three randomly chosen sets of feedback until they reached a high degree of agreement in the different rating categories. They then independently coded all of the feedback provided without consulting each other.

Survey Items

Online Measures.

Motivation to Control Prejudice (Plant & Devine, 1998).

All items are rated on a 1 *(strongly disagree)* to 9 *(strongly agree)*

External Motivation to Control Prejudice Items

1. Because of today's PC (politically correct) standards I try to appear nonprejudiced toward Black people.
2. I try to hide any negative thoughts about Black people in order to avoid negative reactions from others.
3. If I acted prejudiced toward Black people, I would be concerned that others would be angry with me.
4. I attempt to appear nonprejudiced toward Black people in order to avoid disapproval from others.
5. I try to act nonprejudiced toward Black people because of pressure from others.

Internal Motivation to Control Prejudice Items

6. I attempt to act in nonprejudiced ways toward Black people because it is personally important to me.

7. According to my personal values, using stereotypes about Black people is ok. (RC)
8. I am personally motivated to my beliefs to be nonprejudiced toward Black people.
9. Because of my personal values, I believe that using stereotypes about Black people is wrong.
10. Being nonprejudiced toward Black people is important to my self-concept.

Post Feedback Measures.

Perceptions of the Feedback Experience

Please rate your agreement with the following statements:

1. How well do you think your feedback will be received?
   1 *(not well at all)* -7 *(very well)*
2. Are you concerned about how you will be perceived?
   1 *(not at all)* – 7 *(extremely)*

Global Impressions

Please answer the following questions:

1. Please provide a grade from 0 – 100 for the essay you reviewed.
2. How qualified do you believe this student is to receive the scholarship?
   1 *(not at all)* – 7 *(extremely)*
3. To what extent do you think this student should receive the scholarship?
   1 *(definitely not*) – 7 *(Absolutely)*

Self-reported Stereotype Threat

Please rate to what extent you agree with the following statements:
All items rated on a 1 *(strongly disagree)* to 7 *(strongly agree*)

1. When interacting with other races, I worry that people will draw conclusions about my racial group based on my performance
2. I often think about issues concerning race
3. I often feel that people's evaluation of my behavior is based on the racial group that I belong to
4. When interacting with other races, I worry that people will draw conclusions about me based on what they think about my racial group

Self-reported state anxiety and threat emotions

To what extent did you feel the following while giving feedback:
All items rated on a from 1 (*very little*) to 5 (*very much*) scale

1. Tense
2. Under pressure
3. Under strain
4. Nervous/jittery
5. Uneasy

6. Calm (RC)
7. Afraid of not doing well
8. Uncomfortable

Appendix B

Brief Literature Review

Motivation, Stereotypes, and Attitudes. In an effort to be proactive for the chance that the hypotheses are not supported, several measures were created to address possible related, and even confounding, variables. Participant attitudes and motivation to provide effective feedback may impact the results of the proposed study. Past research has found that feedback can differ depending on context and motivation, specifically providing positive feedback is not in the best interest of the student feedback provider, there is less positive feedback (Ho & Yeung, 2014). Due to the fact that this experiment was set in a lab setting, it may lack a context that will motivate the feedback providers to supply accurate and constructive feedback. Thus, motivation to provide constructive feedback was measured to assess a possibly related variable.

In addition to motivation to provide accurate and constructive feedback, a couple other specific measures were included to assess the possible related variables. The author contends that it may be possible for participants' attitudes towards feedback as well as their endorsement of stereotypes regarding STEM students to effect the ability to find significant relationships. Similar to that rationale behind including a motivation to provide accurate and constructive feedback measure, the attitudes toward feedback measure were included to assess possible explanations for null relationships. It may be that some participants do not feel that they can provide accurate feedback or that the participant will not use the feedback. These attitudes were assessed in the case of nonsignificant relationships.

The other original scale included in the proposed study examines stereotypes of STEM students, including stereotypes related to gender. A study conducted by Jones, Ruff, and Paretti (2013) found that men are still seen as having more engineering ability than women. Another study conducted by Schmader, Johns, and Barquissau (2004) found that women in math-related majors that reported higher explicit endorsement of gender stereotypes scored lower on math-related tests when gender stereotypes were salient. The results of these studies indicate that the explicit endorsement of gender stereotypes in STEM may have had an impact on participants' experience during the study.

Measures

Attitudes towards Feedback. A short measure developed by the author was used to assess the participants' attitudes towards feedback. The scale is comprised of six items rated on a 1 *(strongly disagree)* to 5 *(strongly agree)* Likert-type scale. The items include "Providing feedback is crucial for student growth", "Evaluators should provide accurate feedback", "Feedback should be task related", "When providing feedback, evaluators should keep it positive", "Evaluators should give a good effort when providing feedback", and "Negative and positive feedback are equally important to provide".

Motivation to Provide Effective Feedback. The participant's motivation to provide effective feedback was assessed using a short measure rated on a 1 *(strongly disagree)* to 5 *(strongly agree)* Likert-type scale. The items include "I tried to provide useful feedback today", "I gave my full effort when providing feedback

today", "I am motivated to help the candidate today with my feedback", and "I do not care about providing useful feedback to this candidate".

Stereotypes of STEM students. A short measure developed by the author was used to assess the perceptions of the participants of STEM students. The scale is comprised of nine items rated on a 1 *(strongly disagree)* to 5 *(strongly agree)* Likert-type scale with three items specifically addressing gendered STEM stereotypes. These items were derived from common STEM stereotypes. They include "STEM students tend to be intelligent", "STEM students tend to be mentally disciplined", "STEM students tend to be hard working", "STEM students tend to be studious", "STEM students have to work harder than students studying other disciplines", "It is possible that men have more ability in STEMS fields than do women", "In general, men may be better at STEM", and "I don't think that there are any real gender differences in STEM ability (reverse coded)".

Miville-Guzman Universality-Diversity Scale – Short Form, (M-GUDS-S; Pre-Experiment Online Survey) was used to measure diversity perceptions. The scale is comprised of 15 items on a 1 *(Strongly Disagree)* to 7 *(Strongly Agree)* Likert-type scale and is consists of three subscales, including relativistic appreciation (α = .75), senses of connection (α = .72), and diversity of contact (α = .76; Fuertes, Miville, Mohr, Sedlacek, & Gretchen, 2000; Milville et al., 1999).

Cultural Perceptions (Pre-Experimental Online Survey) were measured using a 48 items scale on a 1 *(Strongly Disagree)* to 6 *(Strongly Agree)* Likert-type scale. The scale consists of six subscales aimed at measuring a) Cultural Openness and Desire to Learn (α = .89); (b) Resentment and Cultural Dominance (α = .82); (c) Anxiety and Lack

of Multicultural Self-Efficacy (α =.63); (d) Empathic Perspective-Taking (α =.09); (e) Awareness of Contemporary Racism and Privilege (α =.87); and (f) Empathic Feeling and Acting as an Ally (α =.84; Mallinckrodt, Miles, Bhaskar, Chery, Choi & Sung, 2014; Wang, Davidson, Yakushko, Savoy, Tan & Bleier, 2003).

Appendix C

Analysis Tables for Hypotheses Not Supported

Hypothesis 3 nonsignificant analyses

Table 9

Mediation analysis with student race as IV, four stereotype threat measures as parallel mediators, and negative essay highlights as the DV.

	Consequent (Difference in Negative Highlights)														
	M_1 (Anxiety)			M_2 (Stroop)			M_3 (MAP)			M_4 (Pulse)			Y (Negative Highlights)		
Antecedent	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*
X (FB Race)	-.02	.12	.89	11.17	64.23	.86	9.53	3.71	.01	.15	2.34	.95	-1.49	3.00	.62
M_1													-.65	2.14	.75
M_2													-.00	.00	.26
M_3													-.02	.07	.81
M_4													-.08	.11	.44
Constant	1.95	.19	.00	1066.03	101.55	.00	-8.39	5.87	.15	1.16	3.69	.75	25.95	7.27	.00
	R^2 = .00 $F(1,132)$ = .02, p = .89			R^2 = .00 $F(1,132)$ = .03, p = .86			R^2 = .05 $F(1,132)$ = 6.61, p = .01			R^2 = .00 $F(1,132)$ = .00, p = .95			R^2 = .02 $F(5,128)$ = .58, p = .71		

Table 10

Mediation analysis with student race as IV, four stereotype threat measures as parallel mediators, and Positive FWB (difference between written and spoken comments) as the DV.

	Consequent (Difference in Positive FWB)														
	M_1 (Anxiety)			M_2 (Stroop)			M_3 (MAP)			M_4 (Pulse)			Y (Positive FWB)		
Antecedent	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*
X (FB Race)	-.00	.14	.99	17.84	71.36	.80	10.52	4.27	.02	-.13	2.70	.96	-.35	.38	.36
M_1													-.05	.26	.85
M_2													.00	.00	.81
M_3													-.01	.01	.29
M_4													-.01	.01	.50
Constant	1.95	.22	.00	1047.74	112.83	.00	-10.04	6.76	.14	.67	4.27	.88	-1.90	.93	.04
	R^2 = .00 $F(1,106)$ = .00, p = .99			R^2 = .00 $F(1,106)$ = .06, p = .80			R^2 = .05 $F(1,106)$ = 6.06, p = .02			R^2 = .00 $F(1,106)$ = .00, p = .96			R^2 = .04 $F(5,102)$ = .63, p = .67		

Table 11

Mediation analysis with student race as IV, four stereotype threat measures as parallel mediators, and Negative FWB (difference between written and spoken comments) as the DV.

	Consequent (Difference in Negative FWB)														
	M_1 (Anxiety)			M_2 (Stroop)			M_3 (MAP)			M_4 (Pulse)			Y (Negative FWB)		
Antecedent	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*
X (FB Race)	-.00	.14	.99	17.84	71.36	.80	10.52	4.27	.02	-.13	2.70	.96	-.04	.14	.76
M_1													-.19	.10	.04
M_2													.00	.00	.82
M_3													-.00	.00	.71
M_4													-.00	.01	.89
Constant	1.95	.22	.00	1047.74	112.83	.00	-10.04	6.76	.14	.67	4.27	.88			
	R^2 = .00 F(1,106) = .00, p = .99			R^2 = .00 F(1,106) = .06, p = .80			R^2 = .05 F(1,106) = 6.06, p = .02			R^2 = .00 F(1,106) = .00, p = .96			R^2 = .05 F(5,102) = .97, p = .44		

Table 12

Mediation analysis with student race as IV, four stereotype threat measures as parallel mediators, and Constructive FWB (difference between written and spoken comments) as the DV.

	Consequent (Difference in Constructive FWB)														
	M_1 (Anxiety)			M_2 (Stroop)			M_3 (MAP)			M_4 (Pulse)			Y (Constructive FWB)		
Antecedent	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*
X (FB Race)	-.00	.14	.99	17.84	71.36	.80	10.52	4.27	.02	-.13	2.70	.96	-.38	.51	.45
M_1													.29	.35	.41
M_2													.00	.00	.58
M_3													-.00	.01	.94
M_4													.01	.02	.54
Constant	1.95	.22	.00	1047.74	112.83	.00	-10.04	6.76	.14	.67	4.27	.88	-2.33	1.23	.06
	R^2 = .00 F(1,106) = .00, p = .99			R^2 = .00 F(1,106) = .06, p = .80			R^2 = .05 F(1,106) = 6.06, p = .02			R^2 = .00 F(1,106) = .00, p = .96			R^2 = .02 F(5,102) = .42, p = .84		

Hypothesis 5 nonsignificant analyses

Table 13

Moderated mediation analysis with External MCP as the IV, Internal MCP as the moderator, four stereotype threat measures as parallel mediators, and negative essay highlights as the DV.

		M_1 (Anxiety)			M_2 (Stroop)			M_3 (MAP)			M_4 (Pulse Rate)				Y (Neg. Highlights)		
Antecedent		Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*		Coeff.	*SE*	*p*
X (EMCP)	a_1	-.08	.05	.08	16.97	20.78	.42	.28	1.79	.88	.44	.85	.61	c_1	-1.47	1.28	.26
M_1														b_1	-.10	3.45	.98
M_2														b_2	-.01	.01	.19
M_3														b_3	.01	.10	.91
M_4														b_4			
V (IMCP)	a_2	.06	.04	.17	-44.24	18.42	.02	-.71	1.59	.66	-.89	.76	.24				
IMCP x EMCP	i_1	-.05	.02	.00	-21.23	7.61	.01	-.50	.66	.45	-.39	.31	.22				
Constant		2.91	.35	.00	1838.62	151.72	.00	35.89	13.06	.01	12.89	6.23	.04		41.86	18.14	.02
		R^2 =.26			R^2 = .41			R^2 = .07			R^2 =.09				R^2 =.08		
		$F(5,60)=4.27$, $p=.00$			$F(5,60)=8.41$, $p=.00$			$F(5,60)=.92$, $p=.48$			$F(5,60)=1.22$, $p=.31$				$F(7,58)= .71$, $p=.67$		

N = 66; Control variables: Order of essay and race of student

Table 14

Moderated mediation analysis with External MCP as the IV, Internal MCP as the moderator, four stereotype threat measures as parallel mediators, and Positive FWB (Difference between written and spoken comments) as the DV.

		M_1 (Anxiety)			M_2 (Stroop)			M_3 (MAP)			M_4 (Pulse Rate)				Y (Positive FWB)		
Antecedent		Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*		Coeff.	*SE*	*p*
X (EMCP)	a_1	-.11	.05	.05	23.51	23.72	.33	.07	2.24	.97	.24	.99	.81	c_1	-.15	.19	.44
M_1														b_1	.19	.48	.70
M_2														b_2	.00	.00	.96
M_3														b_3	-.01	.01	.25
M_4														b_4	.03	.03	.22
V (IMCP)	a_2	.10	.05	.04	-19.68	21.97	.38	.22	2.07	.92	-1.29	.91	.16				
IMCP x EMCP	i_1	-.06	.03	.02	-31.32	11.29	.01	-.86	1.06	.42	-.11	.47	.82				
Constant		2.75	.36	.00	1691.93	163.13	.00	24.03	15.37	.12	7.66	6.77	.26		-3.48	3.53	.18
		R^2 =.33			R^2 = .39			R^2 = .03			R^2 =.05				R^2 =.11		
		F(5,47)=4.68, p=.00			F(5,47)=6.11, p=.00			F(5,47)=.30, p=.91			F(5,47)=.52, p=.77				F(7,45)= .78, p=.61		

N = 53; Control variables: Order of essay and race of student

Table 15

Moderated mediation analysis with External MCP as the IV, Internal MCP as the moderator, four stereotype threat measures as parallel mediators, and negative FWB (Difference between written and spoken comments) as the DV.

		M_1 (Anxiety)			M_2 (Stroop)			M_3 (MAP)			M_4 (Pulse Rate)				Y (Negative FWB)		
Antecedent		Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*		Coeff.	*SE*	*p*
X (EMCP)	a_1	-.11	.05	.05	23.51	23.72	.33	.07	2.24	.97	.24	.99	.81	c_1	-.06	.24	.80
M_1														b_1	.24	.61	.70
M_2														b_2	-.00	.00	.50
M_3														b_3	.01	.02	.62
M_4														b_4	.05	.04	.18
V (IMCP)	a_2	.10	.05	.04	-19.68	21.97	.38	.22	2.07	.92	-1.29	.91	.16				
IMCP x EMCP	i_1	-.06	.03	.02	-31.32	11.29	.01	-.86	1.06	.42	-.11	.47	.82				
Constant		2.75	.36	.00	1691.93	163.13	.00	24.03	15.37	.12	7.66	6.77	.26		-1.07	3.21	.74
		R^2 =.33			R^2 = .39			R^2 = .03			R^2 =.05				R^2 =.08		
		$F(5,47)=4.68$, $p=.00$			$F(5,47)=6.11$, $p=.00$			$F(5,47)=.30$, $p=.91$			$F(5,47)=.52$, $p=.77$				$F(7,45)= .53$, $p=.81$		

N = 53; Control variables: Order of essay and race of student

Table 16

Moderated mediation analysis with External MCP as the IV, Internal MCP as the moderator, four stereotype threat measures as parallel mediators, and constructive FWB (Difference between written and spoken comments) as the DV.

		M_1 (Anxiety)			M_2 (Stroop)			M_3 (MAP)			M_4 (Pulse Rate)				Y (Construc. FWB)		
Antecedent		Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*		Coeff.	*SE*	*p*
X (EMCP)	a_1	-.11	.05	.05	23.51	23.72	.33	.07	2.24	.97	.24	.99	.81	c_1	.07	.06	.27
M_1														b_1	-.01	.15	.95
M_2														b_2	.00	.00	.12
M_3														b_3	-.00	.00	.31
M_4														b_4	.02	.01	.02
V (IMCP)	a_2	.10	.05	.04	-19.68	21.97	.38	.22	2.07	.92	-1.29	.91	.16				
IMCP x EMCP	i_1	-.06	.03	.02	-31.32	11.29	.01	-.86	1.06	.42	-.11	.47	.82				
Constant		2.75	.36	.00	1691.93	163.13	.00	24.03	15.37	.12	7.66	6.77	.26		-1.76	.81	.03
		R^2 =.33			R^2 = .39			R^2 = .03			R^2 =.05				R^2 =.24		
		$F(5,47)=4.68$, $p=.00$			$F(5,47)=6.11$, $p=.00$			$F(5,47)=.30$, $p=.91$			$F(5,47)=.52$, $p=.77$				$F(7,45)= 2.07$, $p=.07$		

N = 53; Control variables: Order of essay and race of student

Hypothesis 7 nonsignificant analyses

Table 17

Moderated mediation analysis with NAP as the IV, BOP as the moderator, four stereotype threat measures as parallel mediators, and positive essay highlights as the DV.

		M_1 (Anxiety)			*M_2* (Stroop)			*M_3* (MAP)			*M_4* (Pulse Rate)			Y (Pos. Highlights)		
Antecedent		Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*
X (NAP)	a_1	.43	.27	.12	48.27	124.48	.70	-12.15	9.26	.19	7.04	5.71	.22	9.46	6.71	.16
M_1														1.99	3.21	.54
M_2														-.00	.01	.88
M_3														-.10	.10	.30
M_4														-.37	.15	.02
V (BOP)	a_2	.38	.27	.16	91.15	122.37	.46	7.29	9.11	.43	-3.18	5.62	.57			
NAP x BOP	i_1	.22	.61	.71	-96.59	279.01	.73	20.27	20.76	.33	9.47	12.81	.46			
Constant		2.94	.40	.00	1823.26	183.50	.00	32.74	13.66	.02	8.17	8.42	.34	18.74	17.33	.28
		R^2 =.21 F(5,57)= 3.03, p=.02			R^2 = .28 F(5,57)=4.37, p=.00			R^2 = .08 F(5,57)= 1.02, p=.42			R^2 =.06 F(5,57)= .67, p=.65			R^2 =.17 F(7,55)= 1.66, p=.14		

N = 66; Control variables: Order of essay and race of student

Note: NAP (Implicit Negative Attitudes Toward Prejudice); BOP (Belief One is Prejudiced)

Table 18

Moderated mediation analysis with NAP as the IV, BOP as the moderator, four stereotype threat measures as parallel mediators, and negative essay highlights as the DV.

		M_1 (Anxiety)			M_2 (Stroop)			M_3 (MAP)			M_4 (Pulse Rate)			Y (Neg. Highlights)		
Antecedent		Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*
X (NAP)	a_1	.43	.27	.12	48.27	124.48	.70	-12.15	9.26	.19	7.04	5.71	.22	-.04	7.52	.99
M_1														.72	3.59	.84
M_2														-.01	.01	.15
M_3														.00	.11	.98
M_4														-.20	.17	.25
V (BOP)	a_2	.38	.27	.16	91.15	122.37	.46	7.29	9.11	.43	-3.18	5.62	.57			
NAP x BOP	i_1	.22	.61	.71	-96.59	279.01	.73	20.27	20.76	.33	9.47	12.81	.46			
Constant		2.94	.40	.00	1823.26	183.50	.00	32.74	13.66	.02	8.17	8.42	.34	44.45	19.41	.03
		R^2 =.21 F(5,57)= 3.03, p=.02			R^2 = .28 F(5,57)=4.37, p=.00			R^2 = .08 F(5,57)= 1.02, p=.42			R^2 =.06 F(5,57)= .67, p=.65			R^2 =.07 F(7,55)= .57, p=.77		

N = 66; Control variables: Order of essay and race of student

Note: NAP (Implicit Negative Attitudes Toward Prejudice); BOP (Belief One is Prejudiced)

Table 19

Moderated mediation analysis with NAP as the IV, BOP as the moderator, four stereotype threat measures as parallel mediators, and positive FWB (Difference between written and spoken comments) as the DV.

		M_1 (Anxiety)			M_2 (Stroop)			M_3 (MAP)			M_4 (Pulse Rate)			Y (Positive FWB)		
Antecedent		Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*
X (NAP)	a_1	.32	.29	.28	-23.07	127.13	.86	-12.84	10.45	.23	2.72	6.32	.67	-.33	.94	.73
M_1														.17	.49	.74
M_2														-.00	.00	.71
M_3														-.01	.01	.39
M_4														-.00	.02	.89
V (BOP)	a_2	.23	.29	.44	68.65	127.61	.59	10.46	10.49	.32	-.79	6.35	.90			
NAP x BOP	i_1	.06	.65	.93	-175.14	282.09	.54	9.01	23.18	.39	11.05	14.03	.44			
Constant		2.89	.43	.00	1678.17	189.14	.00	20.28	15.54	.20	6.34	9.41	.50	-2.29	2.70	.40
		R^2 =.18 F(5,57)= 1.98, p=.10			R^2 = .27 F(5,57)=3.40, p=.01			R^2 = .05 F(5,57)= .49, p=.78			R^2 =.04 F(5,57)= .40, p=.85			R^2 =.04 F(7,55)= .27, p=.96		

N = 51; Control variables: Order of essay and race of student

Note: NAP (Implicit Negative Attitudes Toward Prejudice); BOP (Belief One is Prejudiced)

Table 20

Moderated mediation analysis with NAP as the IV, BOP as the moderator, four stereotype threat measures as parallel mediators, and Negative FWB (Difference between written and spoken comments) as the DV.

		M_1 (Anxiety)			M_2 (Stroop)			M_3 (MAP)			M_4 (Pulse Rate)			Y (Negative FWB)		
Antecedent		Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*
X (NAP)	a_1	.32	.29	.28	-23.07	127.13	.86	-12.84	10.45	.23	2.72	6.32	.67	-.30	.32	.35
M_1														-.10	.17	.54
M_2														.00	.00	.13
M_3														-.01	.01	.23
M_4														.01	.01	.52
V (BOP)	a_2	.23	.29	.44	68.65	127.61	.59	10.46	10.49	.32	-.79	6.35	.90			
NAP x BOP	i_1	.06	.65	.93	-175.14	282.09	.54	9.01	23.18	.39	11.05	14.03	.44			
Constant		2.89	.43	.00	1678.17	189.14	.00	20.28	15.54	.20	6.34	9.41	.50	-1.5	.92	.10
		R^2 =.18 F(5,57)= 1.98, p=.10			R^2 = .27 F(5,57)=3.40, p=.01			R^2 = .05 F(5,57)= .49, p=.78			R^2 =.04 F(5,57)= .40, p=.85			R^2 =.18 F(7,55)= 1.33, p=.26		

N = 51; Control variables: Order of essay and race of student

Note: NAP (Implicit Negative Attitudes Toward Prejudice); BOP (Belief One is Prejudiced)

Table 21

Moderated mediation analysis with NAP as the IV, BOP as the moderator, four stereotype threat measures as parallel mediators, and Constructive FWB (Difference between written and spoken comments) as the DV.

		M_1 (Anxiety)			M_2 (Stroop)			M_3 (MAP)			M_4 (Pulse Rate)			Y (Constructive FWB)		
Antecedent		Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*	Coeff.	*SE*	*p*
X (NAP)	a_1	.32	.29	.28	-23.07	127.13	.86	-12.84	10.45	.23	2.72	6.32	.67	1.48	1.10	.18
M_1														.00	.57	.99
M_2														-.00	.00	.36
M_3														.02	.02	.28
M_4														.03	.03	.23
V (BOP)	a_2	.23	.29	.44	68.65	127.61	.59	10.46	10.49	.32	-.79	6.35	.90			
NAP x BOP	i_1	.06	.65	.93	-175.14	282.09	.54	9.01	23.18	.39	11.05	14.03	.44			
Constant		2.89	.43	.00	1678.17	189.14	.00	20.28	15.54	.20	6.34	9.41	.50	-.37	3.14	.91
		R^2 =.18			R^2 = .27			R^2 = .05			R^2 =.04			R^2 =.13		
		F(5,57)=1.98, p=.10			F(5,57)=3.40, p=.01			F(5,57)= .49, p=.78			F(5,57)= .40, p=.85			F(7,55)= .91, p=.51		

N = 51; Control variables: Order of essay and race of student

Note: NAP (Implicit Negative Attitudes Toward Prejudice); BOP (Belief One is Prejudiced)

Table 22
Correlations of Experimental and Control Conditions

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1. IMCP†	-	.24*	-.30*	.22	.07	.14	.08	-.01	-.16	.01	.11	-.08	-.01	.20	-.04	-.17	-.02	-.09	-.07	-.04	-.05	-.04	-.10	-.11
2. EMCP†	.24*	-	-.08	-.02	-.19	.26*	-.14	.02	-.13	.00	-.15	-.25*	.01	-.12	-.10	.01	-.04	.01	.15	-.09	-.02	.15	-.09	.04
3. Race IAT†	-.30*	-.08	-	-.23	-.12	.21	.14	-.05	-.01	-.02	.02	-.12	-.06	-.07	-.01	-.17	-.07	.04	-.04	-.05	-.12	-.11	-.17	.05
4. NAP IAT†	.22	-.02	-.23	-	.27*	.05	.10	.34**	-.08	-.01	.14	.03	.26*	-.19	.05	-.20	.27	.05	.20	.11	.10	.28*	.10	-.21
5. BOP IAT†	.07	-.19	-.12	.27*	-	.03	.16	-.09	-.13	-.02	.04	-.04	.19	-.14	.06	-.39**	.11	-.01	.00	.19	-.01	.02	.11	-.17
6. Stereotype Threat†	.14	.26*	.21	.05	.03	-	.13	.02	.14	.11	-.06	-.24	.21	.03	.15	-.25	-.10	-.12	-.02	-.11	.19	-.04	.02	-.04
7. Anxiety Emotions	.15	-.11	.05	.21	.16	.09	-	.10	-.10	-.15	.14	-.11	.06	.05	-.14	-.07	-.18	.09	.13	-.16	-.03	.17	-.09	.03
8. Stroop Latency	-.14	.09	.05	.06	.02	.15	.35**	-	.20	.06	-.03	-.13	.08	.03	-.16	.02	-.09	.17	.04	-.11	-.01	.00	.32**	-.18
9. MAP	.00	.01	.07	-.11	.16	.15	.16	.31**	-	.16	-.14	-.11	.03	.08	.08	-.01	.07	-.06	-.02	-.06	.16	-.21	.29*	-.17
10. Pulse Rate	-.08	.04	.13	.15	-.04	.28*	.05	.03	.04	-	-.18	-.01	.11	-.07	.27*	-.17	.00	-.14	.14	.06	.20	.00	.26*	-.14
11. Positive Highlights	.20	.06	.06	.13	.02	-.13	.09	-.01	-.16	-.29*	-	.37**	.09	.04	-.19	-.12	-.05	.10	.11	-.10	-.17	.11	.06	.03
12. Negative Highlights	-.02	-.16	-.08	-.04	.01	-.06	.01	-.11	-.01	-.12	.11	-	-.11	.20	.00	.11	-.13	.04	-.12	-.04	-.12	-.17	.13	-.05
13. Positive Comments	.18	.04	.08	.09	.06	.20	-.06	-.06	.12	.05	.30*	-.27*	-	-.42**	.18	-.88**	.24	-.06	.34**	.05	.34**	.39**	.03	-.15
14. Negative Comments	.22	-.07	.01	.02	-.22	-.16	.21	.06	.10	-.03	.16	.26*	.04	-	-.14	.33*	-.73**	.10	-.47**	.01	-.01	-.50**	.07	-.19
15. Constructive Comments	-.12	.01	-.14	-.03	.08	.12	-.07	-.04	.03	-.01	-.22	.42**	-.13	-.11	-	-.05	.21	-.74**	-.19	.10	.71**	-.18	.02	-.01
16. Positive FWB	-.11	-.14	-.08	-.05	.03	-.17	.03	-.01	-.16	.00	-.29*	.28*	-.91**	-.02	.08	-	-.22	.14	-.23	-.20	-.19	-.27	.04	.13
17. Negative FWB	-.39**	.18	-.13	-.10	-.03	.01	-.22	-.03	-.12	-.01	-.23	-.13	-.15	-.74**	.29*	.12	-	-.19	.23	-.05	.12	.17	-.13	.13
18. Constructive FWB	.07	-.07	.15	.17	-.07	-.02	.09	-.05	.06	.23	-.09	-.27	.05	.29*	-.74**	.08	-.37**	-	.11	-.10	-.54**	.05	.09	.06
19. Letter Grade	.13	.14	-.02	.09	.24*	.07	-.13	-.26*	.00	-.12	.20	-.34**	.29*	-.12	-.41**	-.25	-.08	.21	-	-.05	-.11	.66**	-.17	.14
20. Effective Feedback	-.09	.09	.12	-.06	.04	-.01	-.24	-.10	.11	-.11	-.09	-.07	.22	-.27*	-.02	-.18	.09	-.01	.02	-	.05	.07	.02	-.02
21. Video Length	-.20	.07	-.08	.05	.02	.17	.13	.14	.15	.06	-.20	.00	.17	.05	.63**	-.34*	.02	-.45**	-.22	.04	-	-.13	.09	-.19
22. Qualified	.10	.12	.07	-.08	.00	-.12	-.20	-.29*	.03	-.15	.14	-.23	.31*	-.18	-.39**	-.26	.02	.08	.73**	.22	-.37**	-	-.21	.02
23. Race Order	-.10	-.09	-.17	.10	.11	.02	-.33**	-.48**	-.17	-.13	-.04	-.09	.15	-.25*	.12	-.07	.30*	-.12	.32**	.03	.11	.26*	-	.01
24. Essay Order	-.11	.04	.05	-.21	-.17	-.04	-.15	-.17	-.15	-.06	.15	.00	-.07	-.16	.00	.10	.09	.04	-.19	.16	-.06	-.09	.01	-

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

† Variables that were one-time measurements

Experimental relationship located on bottom half of table; Control relationships located on top half of table. Comments refer to video comments.

Table 23

Descriptive Statistics for Independent Variables

Scale	*N*	*M*	*SD*
Internal Motivation to Control Prejudice	67	5.18	1.99
External Motivation to Control Prejudice	67	6.36	1.75
Negative Attitude Toward Prejudice (NAP) IAT	64	-0.61	0.38
Belief Ones Prejudiced (BOP) IAT	68	-0.24	0.38
Race IAT	67	0.40	0.47
STEM Stereotypes	68	4.05	0.57
Attitudes Toward Feedback	67	4.20	0.43
Self-Reported Stereotype Threat	68	2.74	1.32

Table 24

Descriptive Statistics for outcome variables

	Experimental			Control		
	N	*M*	*SD*	*N*	*M*	*SD*
Positive Essay Highlights	67	24.02	16.62	67	19.74	15.61
Negative Essay Highlights	67	17.27	17.55	67	18.98	16.09
Positive Written Comments	54	0.89	1.01	54	0.69	0.85
Constructive Written Comments	54	2.50	2.01	54	2.86	1.99
Negative Written Comments	54	0.58	0.53	54	0.66	0.51
Neutral Written Comments	54	0.22	0.48	54	0.15	0.33
Positive Video Comments	69	3.63	2.26	69	3.01	2.00
Constructive Video Comments	69	4.80	3.03	69	4.77	3.00
Negative Video Comments	69	0.59	0.62	69	0.63	0.76
Neutral Video Comments	69	0.77	0.81	69	0.63	0.71
Positive FWB Comments	54	-2.67	2.16	54	-2.22	1.62
Negative FWB Comments	54	0.00	0.77	54	0.06	0.63
Constructive FWB Comments	54	-2.12	2.65	54	-1.73	2.41
Neutral FWB Comments	54	-0.47	1.02	54	-0.52	0.75
Video Length	69	2.33	1.68	69	2.40	1.49
Letter Grade	69	8.70	2.32	69	8.23	2.04
Number Grade	68	81.63	12.29	68	81.41	7.50
Qualified & Receive Scholarship	69	3.19	0.82	69	2.91	0.69
Desire to Give Effective Feedback	69	4.64	0.41	69	4.56	0.53
Feedback Reception Concern	69	5.36	1.11	69	5.10	1.11
Feedback Perception Concern	69	2.17	1.57	69	1.97	1.36