University of Wisconsin Milwaukee
UWM Digital Commons

Theses and Dissertations

May 2020

Stereotype Threat: Searching for a Double Dissociation of Race-Based Effects and an Explanatory Mechanism

Chandler Jae Zolliecoffer
University of Wisconsin-Milwaukee

Follow this and additional works at: https://dc.uwm.edu/etd

 Part of the Other Psychology Commons, and the Social Psychology Commons

Recommended Citation

Zolliecoffer, Chandler Jae, "Stereotype Threat: Searching for a Double Dissociation of Race-Based Effects and an Explanatory Mechanism" (2020). *Theses and Dissertations*. 2440.
https://dc.uwm.edu/etd/2440

This Thesis is brought to you for free and open access by UWM Digital Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UWM Digital Commons. For more information, please contact open-access@uwm.edu.

STEREOTYPE THREAT: SEARCHING FOR A DOUBLE DISSOCIATION OF RACE-BASED EFFECTS AND AN EXPLANATORY MECHANISM

by

Chandler J. Zolliecoffer

A Thesis Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Master of Science

in Psychology

at

The University of Wisconsin-Milwaukee

May 2020

ABSTRACT

STEREOTYPE THREAT: SEARCHING FOR A DOUBLE DISSOCIATION OF RACE-BASED EFFECTS AND AN EXPLANATORY MECHANISM

by

Chandler J. Zolliecoffer

The University of Wisconsin-Milwaukee, 2020
Under the Supervision of Professor David C. Osmon

Stereotype threat (ST) has been established as a leading theory through which investigators have come to understand and account for discrepancies in performance between stereotyped and non-stereotyped groups. ST has been demonstrated to be a plausible explanation for such discrepant performances favoring White over Black respondents in IQ tests and Black over White respondents in tests of athleticism. The present study was designed to address several gaps in the literature. The current study used ex-Gaussian parameters on traditional simple and two-choice reaction time (RT) tasks of mental 'speed' in place of 'power' measures to address the confound between threat and task difficulty. While results were under-powered, findings suggest interesting but nonsignificant differences in 'speed' tasks warranting further study of the effects in Black IQ and White Athleticism ST. Black IQ-threat comparisons found nonsignificant slowing in RT with fewer attentional lapses while the opposite pattern was found in the White Athletic-threat condition. Additionally, Lexical Decision Task results found partial nonsignificant findings that warrant further study of a thought suppression interpretation of ST effects. Interesting results in the current investigation call for fully powered study of both a double-dissociation between Black-IQ and White-Athleticism ST and the thought suppression mechanism of ST.

© Copyright by Chandler J. Zolliecoffer, 2020
All Rights Reserved

to:

my parents—Loretta & James,

and especially my sister, Jordan;

thank you for your unwavering love and support

TABLE OF CONTENTS

Stereotype Threat: Searching for a Double Dissociation of Race-Based Effects and an Explanatory Mechanism

Over the past two decades, stereotype threat (ST) has proved a successful framework to understand the robust Black/White IQ gap where White test-takers invariably outperform their Black counterparts by nearly one full standard deviation (Nisbett, 2011). Originally identified in pioneering work by Steele and Aaronson (1995), ST theory contends that when members of a stereotyped group are made cognizant of the salient stereotype in evaluative situations, performance suffers. And while ST has been demonstrated to be a plausible explanation for discrepant performance between stereotyped and non-stereotyped groups, there remain many gaps in the literature. The present research addresses several of these gaps.

Notably, less well-examined is the degree to which the ST effect can be observed independent of face-validity. More specifically, the ST phenomenon has chiefly been assessed using cognitively challenging, face-valid 'power' measures. Employing such measures confounds threat and test difficulty, raising the question of whether threat impairs performance only on power tasks viewed inherently as challenging (e.g., IQ tests, the Graduate Record Examination [GRE], mathematics tests). In the present study, traditional simple and two-choice reaction time (RT) tasks or 'speed' tasks are used in place of 'power' measures to evaluate whether the ST effect can be observed on less cognitively challenging tests.

What's more, the ST phenomenon has been demonstrated in non-traditionally stigmatized groups (e.g., White performance on measures of athleticism; Stone et al., 1999). This body of work is far from exhaustive, highlighting yet another gap filled by the present study. Accordingly, in addition to investigating traditional threat effects mentioned above, the present study evaluates a non-traditional athletic-threat effect on 'speed' tasks when such tasks are

presented as a measure of athleticism. Doing so interrogates the expectation of a double-dissociation between the Black and White study participants, thereby extending the ST literature.

Finally, while researchers have begun to make strides toward delineating the mechanism through which ST functions, how such processes work to impair performance requires further exploration. In a review of ST mediators, Pennington and colleagues (2016) identified two classifications of mechanism through which ST may impact performance: affective/subjective and cognitive mechanisms. Recent works lend credibility to the idea that one cognitive mechanism, *thought suppression*, regulates negative emotions and cognitions. It is postulated that thought suppression undermines test performance in situations where ST becomes activated (Schmader & Johns, 2003; Schuster & Schmader, 2015; Schmader et al., 2008). The present study expands upon extant literature, seeking to clarify the mediational role of thought suppression in ST processes on less cognitively-demanding speed tasks. Examining whether thought suppression is associated with speed tasks is hoped to clarify the extent of the ST effect.

A review of the ST literature is provided below, followed by the specific aims and hypotheses of the present study.

A History of Stereotype Threat and Performance

Disparities in performance on psychometric measures of academic and intellectual achievement have long been documented within the field of psychology. Empirical evidence demonstrating support for a phenomenon accounting for such discrepancies, however, was first reported by Steele and Aaronson (1995), coining the term, *stereotype threat*. This foundational study offers an explanation for the discrepant performance observed between members of stigmatized and non-stigmatized groups, particularly within an educational context. The authors defined ST as the phenomenon where members of stigmatized groups are at increased risk of

confirming negative stereotypes when group membership is made salient during evaluative scenarios (Steele & Aaronson, 1995). Steele and Aaronson demonstrated the effect with Black and White Stanford undergraduate students. Results of the investigation found that when a verbal reasoning test was described as diagnostic of intellectual ability, White students significantly outperformed Black students. However, when the same test made no reference to verbal ability, researchers found no significant race-based differences in test performance, and the effect was thus explained as ST. This finding has since been replicated many times over (Brown & Day, 2006; Cohen & Garcia, 2005; MacFarland et al., 2003; Ngyen et al., 2003; Ployhart et al., 2003; Swayer & Hollis-Sawyer, 2005; Smith & Hopkins, 2004).

Replication of the initial effect has led to subsequent exploration into ST theory. Notably, subsequent investigations have evaluated whether the IQ-threat effect could observed with other races and ethnicities. For example, Schmader and Johns (2003) were interested in further elucidating the cognitive processes that underlie ST within a Latino sample. In one of the three conducted experiments, researchers compared the performance of Latino students to that of White students' while under ST conditions. Thirty-three Latino and 40 White participants were randomly assigned to threat or no-threat conditions. Under threat conditions, participants were told that the study task was diagnostic of intellectual ability, whereas those under no-threat were told that the task was a measure of working memory. Results of the study replicated past ST research within this demographic, helping to solidify the robust nature of the ST phenomenon.

Stereotype Threat and Gender. Notably, just as a strong ST effect has been replicated in other minoritized ethnic groups, researchers have found the effect extends also to gender-based performance discrepancies on tests of mathematical ability. In this way, when the stereotype about women being 'bad at math' is made salient to female participants, math

performance suffers, particularly when compared to that of male participants' in the same condition (Spencer et al., 1999; Schmader, 2002). Moreover, the effect is especially pronounced for women who strongly identify with their gender.

Building off of this work, Brown and Pinel (2003) sought to identify mediators and moderators of ST threat on women's mathematics performance while under threat conditions. Results of the study revealed that female participants in high-threat conditions who were *not* highly attuned to, or self-conscious of, their status as a member of a stigmatized group outperformed those who were high in stigma consciousness. This effect was not observed in the low-threat condition such that stigma consciousness was not associated with test performance under these conditions (Brown & Pinel, 2003). Such findings suggest that moderating variables may influence the ways in which ST is or is not experienced by members of stereotyped groups.

Stereotype Threat and Non-Traditional Groups. ST effects are perhaps most readily recognized among traditionally marginalized populations (e.g., individuals from minoritized race/ethnicity or gender backgrounds). Nevertheless, the phenomenon has been observed across a number of other demographic groups suggesting that the effect can generalize more broadly across a number of social categorizations. Most applicable to the current study, Stone and colleagues (1999) endeavored to determine whether non-traditionally stigmatized persons (e.g., White males) experienced performance decrements on a given task as a result of being negatively stereotyped. Thus, the goal of their study was to clarify how the ST mechanism operates within the context of athleticism for Black and White participants. As expected, according the ST theory, the researchers hypothesized that when athletic performance is framed as predictive of "sports intelligence," Black participants performed more poorly than White participants. A double dissociation was also demonstrated in the finding that White participants

underperformed compared to Black participants when the test was framed as a measure of "natural athletic ability." Furthermore, as above in the Brown and Pinel (2003) study, the more strongly White participants identified with athleticism, the greater the ST effect observed. In this way, the authors demonstrate evidence for Steele's original (1997) proposition which argued that ST is a psychological phenomenon that is (1) generalizable and (2) may impact the performance of any individual for whom a stereotype is salient, including the non-traditionally stigmatized. Still, the research support for ST effects observed with non-traditionally stigmatized groups is far from exhaustive. In light of these findings, the current study investigates whether an athletic-threat can be demonstrated on 'speed' tasks described as measures of athleticism. As a result, this study seeks to further expand the ST literature base by testing whether a double dissociation between Black and White participants is evident on 'speed' tasks, un-confounding threat and task difficulty.

Use of Non-Face-Valid Measures to Pinpoint Stereotype Threat

Still, although past research has demonstrated a robust ST effect, one critique that can be leveled against the existing ST literature is its widespread use of face-valid, cognitively challenging 'power' measures to assess the target construct. For example, past research has tended to use subsections of the Wechsler Adults Intelligence Scale (WAIS; Coleman-Carew, 2002), GRE, and mathematical assessments (Wout et al., 2008) and has framed threat in terms of intelligence, even in studies examining athletic ability (Stone et al., 1999). Notably, increased participant anxiety and arousal have been demonstrated to undermine performance across myriad measures, including on tests of mathematical ability (Ashcraft & Kirk, 2001; Hopko et al., 2003), reading comprehension (Everson et., 1994), sporting behaviors, and on computerized neuropsychological tests (Browndyke et al., 2002). It is therefore possible investigators are

capturing a stress-response to tasks perceived as challenging, rather than the ST phenomenon itself (Sommer & Arendasy, 2014; Devine et al., 2012; Johns et al., 2008). That is, because standardized assessments are often cognitively challenging in their own right, a certain level of anticipatory anxiety on the part of the respondent is expected. In this way, fluctuations in performance may reflect anxiety alone, or at the very least in interaction with a ST effect.

As a result, because standardized measures of intelligence and achievement have historically been both face-valid and anxiety-provoking, performance discrepancies may be misattributed to ST. In order, then, to best determine whether the effect observed is reflective of ST, it is appropriate to use non-face-valid, non-power measures of performance to assess the target construct, leaving little room to mischaracterize the observed effect. In light of this, the present study uses traditional simple and traditional two-choice RT tasks as non-face-valid, non-power measures of intelligence and athletic ability.

In his book *Clocking the Mind*, psychologist and author Arthur Jensen details the relationship between cognitive power (intelligence) and cognitive speed (RT) providing compelling evidence for the use of RT speed as a proxy for intelligence (Jensen, 2006). Additionally, this characterization is consistent with the folk wisdom association of intelligence and speed (e.g., 'he is quick to catch on'). By appearance alone, computerized simple and two-choice RT tasks do not seem to be measures of intelligence or athleticism; in this way, any test anxiety that may amount during an assessment on a face-valid measure should not be present on the non-face-valid, cognitive 'speed' tasks. Accordingly, demand characteristics are limited to the cognitive schema elicited by ST task instructions and not anxiety induced by characteristics of the test itself. Given this design, any decrement in performance in the threat condition can more confidently be attributed to ST.

And while the primary aim of the study is to identify whether the ST effect can be observed using non-face-valid 'speed' tasks in place of power tasks, a related, secondary aim of the study endeavors to clarify whether ST is a cognitive schema phenomenon predicated on the saliency of a stereotype. Cognitive schemas guide information-processing, inform how an individual comes to understand the world, and organize generalizations about the self (Oyserman et al., 2003). In this way, repeated exposure to a salient stereotype may begin to negatively impact one's self-evaluation across a broad range of tasks, ultimately hindering performance—even on tasks requiring less cognitive effort. The current study, therefore, assigns Black and White participants to intelligence threat (IQ-threat), athletic-threat, and no-threat control conditions to clarify whether the respective ST effect can be observed on simple and two-choice RT tasks when presented as measures of intelligence and athleticism. If ST activation is demonstrated on non-face-valid 'speed' measures, we might then more confidently conclude that the ST phenomenon is significantly informed by one's cognitive schema, thereby acting independent of task complexity and rigor.

The Role of Thought Suppression in Stereotype Threat

Finally, while research investigating the ST phenomenon is plentiful, the literature offers little consensus on the mechanism through which stereotype threat is activated and for whom it is salient. As described earlier, two classifications of mechanism have been identified in the literature: affective/subjective and cognitive (Pennington et al., 2016). Some of the affective/subjective mediating variables that have been proposed to underpin the effects of ST on performance include: anxiety, individuation tendencies, evaluation apprehension, performance expectancies, and explicit ST endorsement (Pennigton et al., 2016). Alternatively, working

memory, cognitive load, mind wandering, negative thinking, cognitive appraisals, implicit ST endorsement, and thought suppression are some of the proposed *cognitive* mediating variables.

One promising theory suggests that ST is activated by means of *thought suppression.* Those experiencing thought suppression consciously avoid thoughts associated with salient stereotypes while simultaneously, but oftentimes unconsciously, scanning the environment for indications of the stereotype (Logel et al., 2009). In this way, the combination of awareness and avoidance may ultimately lead to the suppression of negative thoughts and emotions that could tax the cognitive resources requisite for adequate performance (Schmader & Johns, 2003; Schuster & Schmader, 2015; Schmader et al., 2008). Because the mental act of suppression draws from the same cognitive resource pool required of individuals to complete the task at hand, this executive reserve becomes quickly depleted. In this way, as a mediator of ST, the thought suppression mechanism proves a plausible explanation for the performance deficits rendered when salient stereotypes are evoked under evaluative scenarios (Schuster et al., 2015; Johns et al., 2008). Depleting cognitive reserve should manifest on both power and speed measures of cognitive ability.

If thought suppression works to activate ST, *post-suppression rebound* is the resultant aftermath of this process. It has been well-documented that subsequent to suppression of intrusive thoughts, individuals experience *post-suppression rebound*—a process whereby previously suppressed thoughts become hyper-accessible (Logel et al., 2009). In their 2009 investigation, Logel and colleagues demonstrate that suppression is the mechanistic force driving the observed decrements in performance for participants under ST by measuring thought suppression both directly and indirectly (post-suppression rebound), and by manipulating how

relevant the stereotype was to the testing situation. We will examine one of the study's experiments more closely as it most clearly relates to the design of the present study.

Female undergraduate students who completed a lexical decision task (LDT) while under math-threat responded slower to threat-related words than to neutral words on a lexical decision task, and slower to threat-related words than male students in the same condition, effectively demonstrating a suppression effect for stereotyped words (Logel et al., 2009). Following release from math-threat, female participants responded faster to threat-related words than to neutral words, and also responded faster to threat-related words than male students in the same condition, effectively demonstrating a post-suppression rebound effect for stereotyped words (Logel et al., 2009). Results of this experiment provide sound empirical evidence to suggest that thought suppression is a viable mechanism that underpins ST.

Therefore, the final aim of the present study is to identify and characterize the ST mechanism using an LDT. It is expected that ST is activated at the individual level for Black participants under IQ-threat and White participants under athletic-threat. As such, the current study uses an initial LDT to evaluate thought suppression while under ST where we would expect slower RTs to threat words, and a second LDT to evaluate post-suppression rebound where we would expect faster RTs to threat words following release from the race-based stereotype manipulation. Understanding the mechanism through which ST operates can offer great insight toward amelioration and even prevention of its effects.

The Present Research: Study Aims and Hypotheses

In accordance with the above literature review, we designed a study to (1) **investigate whether a ST effect could be demonstrated on non-face-valid measures of traditional simple and two-choice RT tasks. These tasks could credibly be framed either as measures**

of intelligence or athleticism in order to activate ST flexibly in either Black or White participants. We therefore predicted that the Black IQ and White Athletic groups would be slower on the ex-Gaussian mu parameter, more variable on the sigma parameter with larger 'fat tails' on the tau parameter than their respective race comparison groups.

Furthermore, as a second aim, the study endeavored to (2) **clarify whether ST is a cognitive schema phenomenon predicated on the saliency of a stereotype.** Accordingly, we hypothesized that the Black IQ group would perform worse than White participants in the same IQ-threat condition on all ex-Gaussian parameters across the simple and two-choice RT tasks. By contrast, we predicted that the White Athletic group would be worse on all ex-Gaussian parameters on the simple and two-choice RT tasks than Black participants in the same athletic-threat condition. Additionally, we expected that there would be no significant differences between Black and White no-threat control groups on all ex-Gaussian parameters on both the simple and choice RT tasks. Finally, we predicted that the racially more salient IQ-threat would show a larger effect than the athletic-threat condition.

Finally, the third aim of the study sought to (3) **characterize the mechanism through which ST is activated.** It has been proposed that thought suppression operates at the time when ST occurs. We therefore hypothesized that the Black IQ group would show a suppression effect on the initial LDT with a slower RT to stereotyped IQ-threat words compared to neutral words. Furthermore, the Black IQ group was predicted to respond slower than the White IQ group to IQ-threat words. Similarly, we expected that the White Athletic group would show a suppression effect on the initial LDT with slower RTs to stereotyped athletic-threat words compared to neutral words. Finally, the White Athletic group was expected to respond slower than the Black Athletic group to athletic-threat words.

Furthermore, we also expected that the Black IQ group and White Athletic group would show a post-suppression rebound effect on stereotyped IQ-threat or athletic-threat words, respectively. Specifically, the Black IQ group and White Athletic groups were expected to respond faster to threat-words than to neutral words on their respective post-suppression LDTs after being released from their respective stereotype manipulation. Moreover, the Black IQ group and White Athletic groups were hypothesized to respond faster to threat words than the White IQ and Black Athletic groups, respectively. Finally, it was predicted that the Black IQ and White Athletic groups would respond significantly faster to their respective threat words on the post-suppression LDTs than to threat words on the suppression LDT, further demonstrating a post-suppression rebound effect.

Method

Participants

Nineteen Black undergraduate students (ages 18–41 years; $M = 24.56$, SD = 6.27) and 32 White (ages 18-42 years; $M = 20.96$, SD = 4.76) from a large, public university in Milwaukee, Wisconsin were recruited for participation in the present study. All participants had completed at least 12 years of education at the time of the study visit. Participants were recruited by flyer and through use of the Sona Experiment Management System, an online study recruitment database for undergraduate students. Both Black and White participants recruited using Sona were awarded extra credit in exchange for their participation in the study. Given the smaller pool of Blacks students from which to recruit within the University, Black students recruited by flyer could elect to receive a $10.00 Amazon gift-card *or* extra credit, in exchange for their participation in the study. Volunteers were excluded from participating in the study if they did not self-identify as Black or White, or any variant thereof (e.g., Caucasian, African-American,

etc.), if they were under 18 years of age, if they were not proficient in English, or if there was a history of psychosis, or neurological insult or injury. Three participants were excluded from data analysis post-consent due to history of psychosis, history of repeated brain injury, and for unreliable completion of the experimental tasks (i.e., rushing through the tasks). All remaining participant data were included in analysis. This included data from 18 Black participants ($n_{IQ\text{-}threat}$ = 6; $n_{athletic\text{-}threat}$ = 6; $n_{control}$ = 6) and 30 White participants ($n_{IQ\text{-}threat}$ = 10; $n_{athletic\text{-}threat}$ = 10; $n_{control}$ = 10). Of the six Black controls, four were randomly assigned to the IQ-threat LDT condition and the remaining two were randomly assigned to the athletic-threat LDT condition. The ten White controls were evenly divided between the IQ-threat and athletic-threat LDT conditions.

Measures

Elementary Cognitive Tasks (ECTs)

The present study uses two ECTs: the traditional simple 0-bit RT task and the 1-bit two-choice RT task. Both ECTs were developed in the Adult Neuropsychology Research Lab and programmed using DirectRT Empirsoft software to control for presentation of stimuli limiting any "noise" that might typically accompany computer processing systems. On the traditional simple ECT, participants were seated in front of a Dell computer and were instructed to press the spacebar on the computer keyboard as soon as the stimulus item (i.e., black dot) appeared on the screen. The stimulus appeared in the center of an otherwise blank screen. Participants were instructed to complete the task as quickly and accurately as possible. Participants completed five practice trials prior to completing 120 real trials. On the two-choice 1-bit ECT, participants were seated in front of a Dell computer and were instructed to make a choice between two options depending on whether the stimulus item appeared on the right or left side of the computer screen (i.e., press the left arrow key if the black dot presented on the left side of the screen, or the right

arrow key if the black dot presented on the right side of the screen). As on the traditional simple 0-bit ECT, participants completed five practice trials prior to completing 120 real trials.

Lexical Decision Tasks (LDTs)

The lexical decision tasks are computerized RT tasks that were used as a measure of thought suppression and post-suppression rebound. The LDTs used in this study were developed in the Adult Neuropsychology Research Laboratory and were programmed using the DirectRT Empirisoft software to control for presentation of stimuli limiting the "noise" that typically accompanies computer processing systems. Participants were seated in front of a Dell computer and instructed to press a designated key to indicate whether the presented stimulus was a real word (i.e., press the right arrow key on the keyboard) or a non-word (i.e., press the left arrow key on the keyboard). The stimulus appeared in the center of the screen in black font across a white background. The presentation of words was randomized. Participants were instructed to complete the task as quickly and accurately as possible. Participants randomized to the IQ-threat or athletic-threat conditions completed one threat-respective suppression LDT immediately following manipulation (i.e., after threat or no-threat statement was read by examiner), and completed another LDT after participants were told that the experimental portion of the study had been completed. Half of the participants in the no-threat control condition completed the IQ-threat suppression and post-suppression LDTs, and the remaining half completed the athletic-threat suppression and post-suppression LDTs. The independent variable of interest was "task type" (i.e., LDT1 vs. LDT2). The dependent variable of interest for both suppression and post-suppression rebound was RT.

The LDTs each comprised 36 letter-strings in total and were split between words and non-words (i.e., 12 threat words, 12 neutral words, and 12 non-words). The IQ-threat LDT

condition included IQ-threat words such as 'ignorant,' 'dishonest,' and 'threatening.' The athletic-threat LDT condition included athletic-threat words such as 'clumsy,' 'slow,' and 'uncoordinated.' Both the IQ-threat and athletic-threat LDT conditions included neutral words such as 'abbreviate,' 'basement,' and 'onwards,' and non-words such as 'dghautre,' 'lededar,' and plotunuium,' Orthographic frequencies and number of letters within each letter-string were obtained for each category using MCWord: An Orthographic Wordform Database (Medler, 2005). Mean orthographic frequency and mean number of letters within each letter-string were also calculated. See the Appendix for a complete list of letter-strings included within the LDTs and for mean frequency and letter-string length values for each category. The independent variable of interest was word type (i.e., threat, neutral, or non-word). Because the LDTs used in the study were sufficiently easy, RT, as opposed to accuracy, served as the dependent variable and was used as a measure of both suppression and post-suppression rebound effects.

Assignment to Experimental and Control Conditions

Prior to their arrival, participants were assigned to the intelligence-threat (IQ-threat), athletic-threat or the control (no-threat control) condition using modified random assignment in order to preserve statistical efficiency and power. Furthermore, participants in both threat conditions were then assigned to their respective IQ-threat or athletic-threat suppression and post-suppression LDT conditions. Half of the participants in the no-threat control condition were randomly assigned to the IQ-threat suppression and post-suppression LDTs, and the other half were randomly assigned to the athletic-threat suppression and post-suppression LDTs.

Procedure

Participants were run individually. Upon arrival, participants were consented by a Hispanic female examiner. During the consent process, participants were told that they would be

completing several computerized tasks in addition to some paper and pencil tasks that would "ultimately help us study the nature of reaction time and also help us gain a better understanding of the construct validity of elementary cognitive tasks." After being consented, participants answered demographic questions and provided information about their age, sex, education, race, primary language spoken, handedness, psychiatric disorder history, learning disability history, neurologic disorder history, head trauma history, current prescription medications, and vision-related problems.

After collecting this demographic information and background data, the examiner explained to participants that the purpose of this study was to "catalogue the range of scores on measures of reaction time frequently used in psychological research." Those in the no-threat control condition were then told that their participation in this research would "help develop normative values for a typical sample of adults on common reaction time tasks that are used in psychology." And while participants in the IQ- and athletic-threat conditions were also told that their participation in this research would "help develop normative values for a typical sample of adults on common reaction time tasks," instructions diverged by threat-condition after this initial instruction. In this way, those in the IQ-threat condition were then told by the examiner that the tasks they would complete over the course of the study session were "genuine tests of neural speed which is the basis of intelligence" whereas those in the athletic-threat condition were told that the tasks they would complete over the course of the study session were "genuine tests of reaction time speed which is the basis of athletic ability."

Following manipulation, participants completed the first LDT and then subsequently completed the 0-bit and 1-bit ECTs, respectively. After completion of the simple and two-choice ECTs, the examiner read the following statement to the participants so as to release them from

suppression: “the experimental portion of the study is now complete. The remaining tasks are to help us better understand you as a person.” Participants then completed the second, condition-specific LDT.

Finally, following completion of all study-related tasks, the examiner debriefed the participants and informed them of the true purpose of the study, thanked them for their participation, and asked that participants not discuss the study or its true purpose with anyone else in order to preserve the integrity of the study.

Power Analysis

In their meta-analysis of ST literature, Nguyen & Ryan (2008) establish that moderately-explicit ST-activating cues consistently produce a moderate effect, mean d = .|64|, for minoritized individuals. As such, the equivalent f statistic of .321, was used in the calculation of an *a priori* power analysis. Results of the power analysis reveal a sample size requirement of N = 131. Accordingly, we appreciate the fact that the acquired sample for the present investigation was underpowered. Results of the study should therefore be interpreted with extreme caution.

Data Processing and Analysis

Past literature suggests that RT distributions are generally positively skewed (Luce, 1986); this holds true in the present study as well. As such, before conducting all RT analyses, data were trimmed of outliers to minimize invalid RTs. To do this, the Quantile Range Outliers method, which is a conservative trimming procedure, was employed using the JMP v.14 data analysis software (JMP, 2018).

Furthermore, the ex-Gaussian distribution is commonly found to fit RT distributions well, although other distributions might fit better (Luce, 1986). As it pertains to the present study, the empirical data were determined to be non-normally distributed. As such, the Aikaike’s

Information Criterion-corrected identified the ex-Gaussian model to be the best distributional fit of the empirical RT data. Accordingly, the mu, sigma, and tau parameters of this distribution were used to compare results between groups. And while it was originally proposed that Wilcoxon signed-rank tests would be performed to test group differences on ECTs, because results did not differ between parametric and nonparametric analysis methods, the more familiar parametric results are reported below for clarity and simplicity.

Results

ECT Performance

Aim 1

In the first aim of the study, we proposed to investigate whether a ST effect could be observed on non-face-valid measures of intelligence and athleticism through assessment of RT parameters on simple and two-choice RT tasks.

Hypothesis 1

The Black IQ and White Athletic groups were expected to be slower on the ex-Gaussian mu parameter, more variable on the sigma parameter with larger 'fat tails' on the tau parameter than their respective same-race control groups. These results were expected due to the effect of ST, which has been shown to disrupt performance in 'power' IQ tests. Thus, the ST phenomenon was expected to affect 'speed' tests similarly on both the simple and two-choice RT ECTs.

Simple RT Mu. Results of a one-way between-subjects ANOVA and subsequent post-hoc test demonstrated no statistically significant differences between the Black IQ and Black Control groups $F(2, 15) = 1.14$, $p = 0.35$, although the Black IQ group was non-significantly slower ($M_{\mu} = 275.88$, $SD_{\mu} = 37.87$) than the control group ($M_{\mu} = 254.25$, $SD_{\mu} = 35.57$). Similarly, results of an ANOVA revealed no statistically significant differences between the White Athletic and White Control groups, $F(2, 26) = 0.55$, $p = 0.58$. In fact, results were in the

opposite direction from prediction such that the White Athletic group was faster (M_{μ}= 256.05, SD_{μ} = 22.90) than the White Control group (M_{μ} = 271.29, SD_{μ}= 53.20). Interestingly, the White IQ group was slowest (M_{μ} = 309.31, SD_{σ} = 56.43).

Choice RT Mu. Results of a one-way between-subjects ANOVA and subsequent post-hoc test demonstrated no statistically significant differences between the Black IQ and Black Control groups $F(2, 15) = 1.99$, $p = 0.17$, although, again, the Black IQ group was slower (M_{μ} = 293.65, SD_{μ} = 27.17) than the Black Control group (M_{μ} = 272.66, SD_{μ} =19.48). Curiously, the Black Athletic group was slowest (M_{μ} = 322.10, SD_{μ} = 66.62). Similarly, results of an ANOVA and post-hoc test revealed no statistically significant differences between the White Athletic and White Control groups $F(2, 26) = 0.40$, $p = 0.67$. In fact, results were in the opposite direction from prediction such that the White Athletic group was faster (M_{μ} = 280.49, SD_{μ} = 21.18) than the White Control group (M_{μ} = 283.91, SD_{μ} = 31.25). The White IQ group was fastest (M_{μ} = 272.69, SD_{μ} = 28.72).

Simple RT Sigma. Results of a one-way between-subjects ANOVA and subsequent post-hoc test demonstrated no statistically significant differences between the Black IQ and Black Control groups $F(2, 15) = 1.56$, $p = 0.34$, although the Black IQ group was more variable (M_{σ} = 26.82, SD_{σ} =14.70) than the Control group (M_{σ} = 19.97, SD_{σ} = 6.90), as was the Black Athletic group (M_{σ} = 28.46, SD_{σ} = 7.20). Similarly, results of an ANOVA revealed no statistically differences between the White Athletic and White Control groups $F(2, 15) = 0.10$, $p = 0.90$. In fact, results were in the opposite direction from prediction such that the White Athletic group was less variable (M_{σ} = 18.97, SD_{σ} = 4.96) than the White Control group (M_{σ} = 27.17, SD_{σ} = 21.50). The White IQ group was less variable (M_{σ} = 20.51, SD_{σ} = 6.62) than controls.

Choice RT Sigma. Results of a one-way between-subjects ANOVA and subsequent post-hoc test demonstrated no statistically significant differences between the Black IQ and Black Control groups, $F(2, 25) = 0.60$, $p = 0.56$, although, again, the Black IQ group was more variable ($M_\sigma = 31.47$, $SD_\sigma = 8.33$) than the Black Control group ($M_\sigma = 25.78$, $SD_\sigma = 8.52$). Interestingly, the Black Athletic group was also more variable than controls ($M_\sigma = 30.82$, $SD_\sigma = 12.26$). Similarly, results of an ANOVA and post-hoc test revealed no statistically significant differences between the White Athletic and White Control groups $F(2, 26) = 0.84$, $p = 0.44$. In fact, results were in the opposite direction from prediction such that the White Athletic group was less variable ($M_\sigma = 29.91$, $SD_\sigma = 7.45$) than the White Control group ($M_\sigma = 31.64$, $SD_\sigma = 9.44$). The White IQ group was least variable ($M_\sigma = 26.99$, $SD_\sigma = 5.82$).

Simple RT Tau. Results of a one-way between-subjects ANOVA and subsequent post-hoc test demonstrated no statistically significant differences between the Black IQ and Black Control groups $F(2, 15) = 2.38$, $p = 0.13$, although the Black IQ group had fewer responses in the 'fat tail' ($M_\tau = 39.09$, $SD_\tau = 24.72$) than the Control group ($M_\tau = 61.14$, $SD_\tau = 31.02$), while the Black Athletic group had more 'fat tail' responses ($M_\tau = 76.53$, $SD_\tau = 33.16$). Similarly, results of an ANOVA and post-hoc test demonstrated no statistically significant differences between the White Athletic and White Control groups $F(2, 15) = 0.72$, $p = 0.498$, although the White Athletic group had more responses in the 'fat tail' ($M_\tau = 51.61$, $SD_\tau = 25.51$) compared to the White Control group ($M_\tau = 42.55$, $SD_\tau = 20.10$). The White IQ group also had more 'fat tail' responses ($M_\tau = 52.86$, $SD_\sigma = 14.77$) than controls.

Choice RT Tau. Results of a one-way between-subjects ANOVA and subsequent post-hoc test demonstrated no statistically significant differences between the Black IQ and Black Control groups $F(2, 25) = 1.28$, $p = 0.31$, and again, the Black IQ group had fewer 'fat tail'

responses ($M_\tau = 46.92$, $SD_\tau = 20.84$) than the Black Control group ($M_\tau = 60.37$, $SD_\tau = 23.28$). The Black Athletic group had the fewest 'fat tail' responses ($M_\tau = 41.37$, $SD_\tau = 19.12$). Similarly, results of a one-way ANOVA and post-hoc test revealed no statistically significant differences between the White Athletic and White Control groups $F(2, 26) = 1.27$, $p = 0.299$, although results were in the predicted direction such that the White Athletic group had more 'fat tail' responses ($M_\tau = 53.81$, $SD_\tau = 17.11$) than the White Control group ($M_\tau = 41.83$, $SD_\tau = 24.04$). The White IQ group was almost identical to the controls ($M_\tau = 41.70$, $SD_\tau = 16.65$).

ST Effect Differences

Aim 2

In the second aim of the study, we sought to clarify whether the ST effect is a cognitive schema phenomenon predicated on the saliency of a stereotype.

Hypothesis 2

It was expected that Black IQ group would perform worse on all ex-Gaussian parameters than White participants in the IQ-threat condition on both the simple and choice RT tasks. Conversely, the White Athletic group was expected to be worse on all ex-Gaussian parameters on both the simple and choice RT tasks than Black participants in the same athletic-threat condition. Further, it was expected that there would be no significant differences between Black and White no-threat control groups on all ex-Gaussian parameters on both the simple and choice RT tasks. Finally, it was expected that the longer and more racially salient stereotype of the IQ-threat condition would show a larger effect than the athletic-threat condition.

Simple RT Mu. Results of a planned contrast in a 2 (race: Black vs. White) X 3 (condition: IQ- vs. No- vs. Athletic-threat) two-way ANOVA showed no statistically significant differences between the Black IQ and White IQ groups $F(1, 41) = 0.01$, $p = 0.94$, and the two

groups were nearly identical (M_{μ} = 275.88, SD_{μ} = 37.66 vs. M_{μ} = 277.61, SD_{μ} = 56.43, respectively). Similarly, results of an ANOVA planned contrast revealed no statistically differences between the White Athletic and Black Athletic groups $F(1, 41) = 1.90$, $p = 0.18$. In fact, results were in the opposite direction from prediction with the White Athletic group being faster (M_{μ} = 256.05, SD_{μ} = 22.90) than the Black Athletic group (M_{μ} = 286.99, SD_{μ} = 40.989). Additionally, the planned contrast between the Black and White no-threat control condition demonstrated no statistically significant differences $F(1, 41) = 0.58$, $p = 0.45$. Finally, effect sizes were not examined due to the lack of significant differences between racial groups across threat conditions.

Choice RT Mu. Results of a planned contrast in a 2 (race: Black vs. White) X 3 (condition: IQ- vs. No- vs. Athletic-threat) ANOVA showed no statistically significant differences between the Black IQ and White IQ groups $F(1, 41) = 1.37$, $p = 0.249$, though the Black IQ group was non-significantly slower (M_{μ} = 293.65, SD_{μ} = 27.17) than the White IQ group ($M\mu$ = 272.69, SD_{σ} = 28.72), as predicted. By contrast, results of the ANOVA planned contrast revealed a statistically significant difference between the White Athletic and Black Athletic groups $F(1, 41) = 5.62$, $p = 0.023$; however, results were in the opposite direction from prediction such that the White Athletic group was faster (M_{μ} = 280.49, SD_{σ} = 21.18) than the Black Athletic group (M_{μ} = 322.10, SD_{σ} = 66.62). The effect sizes for this comparison is delta = 11.88 yielding only an adjusted power of .53 at $\alpha = .05$, although the confidence interval for this effect size does not include zero. Additionally, the planned contrast between the Black and White no-threat control condition demonstrated no statistically significant differences, $F(1, 41) = 0.39$, $p = 0.53$.

Simple RT Sigma. Results of a planned contrast in a 2 (race: Black vs. White) X 3 (condition: IQ- vs. No- vs. Athletic-threat) ANOVA showed no statistically significant differences between the Black IQ and White IQ groups $F(1, 41) = 0.93$, $p = 0.34$, though the Black IQ group ($M_\sigma = 26.82$, $SD_\sigma = 14.70$) was more variable than the White IQ group ($M_\sigma = 20.51$, $SD_\sigma = 6.62$), as predicted. Similarly, results of an ANOVA planned contrast revealed no statistically significant differences between the White Athletic and Black Athletic groups $F(1, 41) = 2.19$, $p = 0.15$. In fact, results were in the opposite direction from prediction such that the White Athletic group was less variable ($M_\sigma = 18.97$, $SD_\sigma = 4.96$) than the Black Athletic group ($M_\sigma = 28.46$, $SD_\sigma = 7.20$). Additionally, the planned contrast between the Black and White no-threat control condition demonstrated no statistically significant differences $F(1, 41) = 1.26$, $p = 0.27$. Finally, effect sizes were not examined due to the lack of significant differences.

Choice RT Sigma. Results of a planned contrast in a 2 (race: Black vs. White) X 3 (condition: IQ- vs. No- vs. Athletic- threat) ANOVA showed no statistically significant differences between the Black IQ and White IQ groups $F(1, 41) = 983$, $p = 0.33$, and the Black IQ group was more variable ($M_\sigma = 31.47$, $SD_\sigma = 8.33$) than the White IQ group ($M_\sigma = 26.99$, $SD_\sigma = 5.82$). Similarly, results of the ANOVA contrast revealed no statistically significant differences between the White Athletic and Black Athletic groups, $F(1, 41) = 0.42$, $p = 0.84$. In fact, results were in the opposite direction from prediction such that the White Athletic group was less variable ($M_\sigma = 29.91$, $SD_\sigma = 7.45$) than the Black Athletic group ($M_\sigma = 30.82$, $SD_\sigma = 12.26$). Additionally, the planned contrast between the Black and White no-threat control condition demonstrated no statistically significant differences $F(1, 41) = 1.68$, $p = 0.20$. Finally, effect sizes were not examined due to the lack of significant differences.

Simple RT Tau. Results of a planned contrast in a 2 (race: Black vs.White) X 3 (IQ- vs. No- vs. Athletic- threat) ANOVA showed no statistically significant differences between the Black IQ and White IQ groups $F(1, 41) = 1.37$, $p = 0.29$, and the Black IQ group had fewer responses in the 'fat tail' than the White IQ group ($M_\tau = 39.09$, $SD_\tau = 24.72$ vs. $M_\tau = 52.86$, $SD_\tau = 14.77$, respectively), contrary to prediction. Similarly, results of an ANOVA contrast revealed no statistically significant differences between the White Athletic and Black Athletic groups $F(1, 41) = 3.88$, $p = 0.056$, and the White Athletic group had fewer responses in the 'fat tail' ($M_\tau = 51.60$, $SD_\tau = 25.51$) compared to the Black Athletic group ($M_\tau = 76.53$, $SD_\tau = 33.16$), contrary to prediction. Additionally, the planned contrast between the Black and White no-threat control condition demonstrated no statistically significant differences $F(1, 41) = 2.16$, $p = 0.149$. Finally, effect sizes were not examined due to the lack of significant differences.

Choice RT Tau. Results of a planned contrast in a 2 (race: Black vs. White) X 3 (IQ- vs. No- vs. Athletic-threat) ANOVA showed no statistically significant differences between the Black IQ and Black Control groups $F(1, 41) = 025$, $p = 0.617$, and the Black IQ group had more 'fat tail' responses ($M_\tau = 46.92$, $SD_\tau = 20.84$) than the White IQ group ($M_\tau = 41.70$, $SD_\tau = 16.65$). Similarly, results of the ANOVA contrast demonstrated no statistically differences between the White Athletic and Black Athletic groups $F(1, 41) = 0.26$, $p = 0.62$, although results were in the predicted direction with the White Athletic having more 'fat tail' responses ($M_\tau = 53.81$, $SD_\tau = 17.11$) compared to the Black Athletic group ($M_\tau = 41.37$, $SD_\tau = 19.12$). Also, the planned contrast between Black and White no-threat control condition demonstrated no statistically significant differences $F(1, 41) = 3.08$, $p = 0.09$. Finally, effect sizes were not examined due to the lack of significant differences.

Suppression and Post-Suppression Rebound

Aim 3

In the third aim of the study, we identified and characterized the mechanism through which the ST effect is activated at the individual level. Again, Wilcoxon-signed rank tests were initially proposed to test group differences across word type, however, because results did not differ between parametric and nonparametric methods, the more familiar parametric results are reported here. We analyzed participant reaction times to stereotyped words and neutral words in a four-way factorial 2 (race: Black vs. White) X 3 (condition: IQ- vs. No- vs. Athletic-threat) X 2 (task: LDT1 vs. LDT 2) X 3 (word type: IQ threat word vs. neutral word vs. athletic threat word) ANOVA. Though the four-way interaction was not significant, $F(3, 2403) = 1.46$, $p = 0.22$, the omnibus F was significant, $F(43, 2403) = 5.68$, $p < 0.0001$, prompting use of planned contrasts to perform follow-up post-hoc tests.

Hypothesis 3a

It was predicted that the Black IQ group would show a suppression effect of stereotyped IQ-threat words (i.e., slower RT). In this way, the Black IQ group was expected to respond slower to IQ-threat words than to neutral words on the suppression LDT. Furthermore, the Black IQ group was expected to respond slower to IQ-threat words than the White IQ group. Similarly, it was predicted that the White Athletic group would show a suppression effect of stereotyped athletic-threat words (i.e., slower RT). In this way, the White Athletic group was expected to respond slower to athletic-threat words than to neutral words on the suppression LDT. Furthermore, the White Athletic group was expected to respond slower to athletic-threat words than the Black Athletic group.

Black IQ Suppression. Results of a planned contrast in a four-way factorial 2 (race: Black vs. White) X 3 (condition: IQ- vs. No- vs. Athletic-threat) X 2 (task: LDT1 vs. LDT 2) X 3 (word type: IQ threat word vs. neutral word vs. athletic threat word) ANOVA demonstrated no statistically significant differences in RT between IQ-threat words and neutral words among Black participants on the suppression LDT, $F(3, 2403) = 0.46$, $p = 0.4992$. In fact, results were in the opposite direction from prediction such that the Black IQ group responded faster to IQ-threat words (M = 810.43, S*D* = 267.75) than to neutral words (M = 860.93, SD = 371.44). Effect sizes were not examined due to lack of significance. Results of a planned contrast did reveal statistically significant differences in RT to IQ-threat words when comparing Black and White IQ groups on the suppression LDT, $F(1, 2403) = 3.93$, $p = 0.048$. However, results were in the opposite direction from prediction such that the Black IQ group responded faster to IQ-threat words (M = 810.43, SD = 267.75) than the White IQ group (M = 941.26, SD = 297.42), and the effect, $d = 0.00$, was not significant as confidence interval includes zero.

White Athletic Suppression. Results of a planned contrast in a four-way factorial 2 (race: Black vs. White) X 3 (condition: IQ- vs. No- vs. Athletic-threat) X 2 (task: LDT1 vs. LDT 2) X 3 (word type: IQ threat word vs. neutral word vs. athletic threat word) ANOVA demonstrated a statistically significant difference in RT between athletic-threat words and neutral words among White participants on the suppression LDT, $F(1, 2403) = 4.03$, $p = 0.045$. Though the results were significant, they were not significant in the predicated direction such that the White Athletic group responded faster to athletic-threat words (M = 881.67, SD = 329.26) than to neutral words (M = 993.25, SD = 416.67). Results of the planned contrast yielded a small effect, $d = 0.25$ at alpha = .05. Results of a separate planned contrast also demonstrated statistically significant differences in RT to athletic-threat words when comparing

the White and Black Athletic groups $F(1, 2403) = 6.22$, $p = 0.013$. Results, again however, were not in the predicted direction such that the White Athletic group responded faster to athletic-threat words (M = 881.67, SD = 329.26) than the Black Athletic group (M =1058.82, SD = 558.75). Results of the planned contrast yielded a small effect, $d = 0.43$, although the confidence interval for this effect size includes zero.

Hypothesis 3b

It was predicated that the Black IQ group would show a post-suppression rebound effect of stereotyped IQ-threat words (i.e., faster RT) after being released from suppression. In this way, the Black IQ group was expected to respond faster to IQ-threat words than to neutral words on the post-suppression LDT. Furthermore, the Black IQ group was expected to respond faster to IQ-threat words than the White IQ group.

Similarly, it was predicted that the White Athletic group would show a post-suppression rebound effect of stereotyped athletic-threat words (i.e., slower RT) after being released from suppression. In this way, the White Athletic group was expected to respond faster to athletic-threat words than to neutral words on the post-suppression LDT. Furthermore, the White Athletic group was expected to respond faster to athletic-threat words than the Black Athletic group.

Finally, it was predicted that the Black IQ group would respond significantly faster to IQ-threat words on the post-suppression LDT than to IQ-threat words on the suppression LDT further demonstrating a post-suppression rebound effect. Similarly, the White Athletic group was predicted to respond significantly faster to athletic-threat words on the post-suppression LDT than to athletic-threat words on the suppression LDT, further demonstrating a post-suppression rebound effect.

Black IQ Post-Suppression Rebound. Results of a planned contrast in a four-way factorial 2 (race: Black vs. White) X 3 (condition: IQ- vs. No- vs. Athletic-threat) X 2 (task: LDT1 vs. LDT 2) X 3 (word type: IQ threat word vs. neutral word vs. athletic threat word) ANOVA demonstrated no statistically significant differences in RT between IQ-threat words and neutral words among Black participants on the post-suppression LDT, $F(1, 2403) = 2.52$, $p = 0.113$. However, results were in the predicted direction such that the Black IQ group responded faster to IQ-threat words (M = 703.43, SD = 167.39) than to neutral words (M = 822.02, SD = 329.08). Effect sizes were not examined, however, due to lack of significance. Additionally, results of a separate planned contrast revealed statistically significant differences in the predicted direction, when comparing the Black and White IQ groups on the post-suppression LDT, $F(1, 2403) = 4.24$, $p = 0.04$, such that the Black IQ group responded faster to IQ-threat words (M = 703.43, SD = 167.39) than the White IQ group (M = 839.14, SD = 345.02). Results of the planned contrast, however, did not yield a significant effect, $d = 0.00$, as the confidence interval includes zero.

White Athletic Post-Suppression Rebound. Results of a planned contrast in a four-way factorial 2 (race: Black vs. White) X 3 (condition: IQ- vs. No- vs. Athletic-threat) X 2 (task: LDT1 vs. LDT 2) X 3 (word type: IQ threat word vs. neutral word vs. athletic threat word) ANOVA demonstrated statistically significant differences in RT between athletic-threat words and neutral words among White participants on the post-suppression LDT, $F(1, 2403) = 6.69$, $p = 0.01$. Results were in the predicted direction such that the White Athletic group responded faster to athletic-threat words (M = 814.15, SD = 241.64) than to neutral words (M = 944.47, SD = 359.62). Results of the planned contrast, however, did not yield a significant effect, $d = 0.00$, as the confidence interval includes zero. Additionally, results of a separate planned contrast

revealed statistically significant RT differences to athletic-threat words when comparing White and Black participants on the post-suppression LDT, $F(1, 2403) = 6.22$, $p = 0.013$, such that the White Athletic group responded faster (M = 814.15, SD = 241.64) to athletic-threat words than the Black Athletic Group (M = 973.12, SD = 471.39). Results of the planned contrast, however, did not yield a significant effect, $d = 0.00$, as the confidence interval includes zero.

Black IQ Suppression vs. Post-Suppression LDT. Results of a planned contrast in a four-way factorial 2 (race: Black vs. White) X 3 (condition: IQ- vs. No- vs. Athletic-threat) X 2 (task: LDT1 vs. LDT 2) X 3 (word type: IQ threat word vs. neutral word vs. athletic threat word) ANOVA revealed no statistically significant RT differences between IQ-threat words among Black participants on the suppression LDT and post-suppression LDT, $F(1, 2403) = 2.05$, $p = 0.152$. Results, however, were in the predicted direction such that the Black IQ group responded faster on the post-suppression LDT (M = 703.43, SD = 167.39) and slower on the suppression LDT (M = 810.43, SD = 267.75). Effect sizes were not examined due to lack of significance.

White Athletic Suppression vs. Post-Suppression LDT. Results of a planned contrast in a four-way factorial 2 (race: Black vs. White) X 3 (condition: IQ- vs. No- vs. Athletic-threat) X 2 (task: LDT1 vs. LDT 2) X 3 (word type: IQ threat word vs. neutral word vs. athletic threat word) ANOVA revealed statistically significant RT differences between athletic-threat words among White participants on the suppression LDT and post-suppression LDT, $F(1, 2403) = 6.22$, $p = 0.013$. Results were in the predicted direction such that the White Athletic group responded faster on the post-suppression LDT (M = 814.15, SD = 241.64) and slower on the suppression LDT (M = 881.67, SD = 329.26). Moreover, results of the planned contrast yielded a large effect, $d = 2.40$, at $\alpha = .05$.

Discussion

Overall, results of the investigation were largely not confirmed by the study's hypotheses; this is likely attributed to insufficient power given the smaller sample size. The ex-Gaussian and LDT results did, however, show interesting trends. More specifically, study findings evidenced support for a post-suppression rebound effect partially confirming the third hypothesis. As such, it is clear that some results deserve further consideration, and warrant continued data collection to better interrogate the study's hypotheses. This will be further explicated throughout this section.

The first aim of the study sought to establish whether a ST effect could be observed on non-face-valid measures of intelligence and athleticism on simple and choice ECTs. It was predicted that stigmatized groups under their respective ST conditions would perform worse on RT parameters across simple and choice ECTs. And while Black participants under ST were slower (mu) and more variable (sigma) on both simple and choice RT tasks compared to their respective same-race controls, the groups were not significantly different. By contrast, the Black IQ group had fewer RTs in the 'fat tail' (tau) of the RT task distributions compared to their controls on both simple and choice RT tasks, contrary to prediction. If upheld by results with sufficient power to detect an effect, these findings would suggest that ST does alter performance for Black participants, even on non-face-valid tasks. That is to say, if these results hold up under further scrutiny, then ST would appear to be a cognitively disruptive phenomenon on efficient RT (mu) that is not dependent upon the difficulty of the task. It is unclear why the tau results did not follow prediction, but it can be speculated that ST focuses attention. Such a focusing effect may reduce attentional lapses that result in very long RTs and a 'fatter tail' on 'speed' tasks.

Moreover, White participants under ST showed an unusual inverse pattern of results compared to Black participants under IQ threat. Specifically, White participants under athletic-threat were faster and less variable than their controls, and had more inefficient RTs in the 'fat tail' on both simple and choice RT tasks. These findings are contrary to the results observed by Black participants under IQ-threat and are difficult to interpret within the framework of ST theory. The results seem to suggest that 'speed' tasks under athletic-threat do not behave how 'power' tests behave, or even how 'speed' tasks behave under IQ-threat. It might be speculated that White participants under threat marshal resources, improving speed and variability at the expense of greater lapses in attention, producing a 'fatter' tail. Further research with sufficient sample power is warranted to follow-up on the suggestions from aim one results.

The second aim of the study sought to clarify whether the ST effect is consistent with a cognitive schema phenomenon predicated on the salient stereotype. Accordingly, we expected again to see a stigmatized group ST effect with worse performance on RT parameters across simple and choice ECTs. Results comparing Black and White participants generally support this interpretation, however, they further suggest that task complexity moderates results when considering racial group, as discussed next.

Specifically, the inverse pattern of performance between Black and White participants discussed above on this study's first aim also generally holds for the stigmatized groups analysis in the second aim. However, it holds better for the more complex choice RT task compared to the easier simple RT task. These findings suggest that task complexity acts as a moderator of the stigmatized groups' performance. Future exploration into the moderating factors of ST performance is therefore warranted, especially in light of the literature that finds ST effects tend to hold better for individuals who identify more with the salient stereotype (e.g., Smith &

Hopkins, 2004, Davis et al., 2006). That is, stigmatized group identification may be clouding the results, resulting in individual differences that moderate group results. It is possible that different individuals within a group have opposite patterns of results in the RT parameters, such that group results do not reflect individual patterns of performance for a significant portion of the group. Only studies with greater power than the current study will be able to sort out these issues.

Finally, the third aim of the study sought to identify and characterize the mechanism through which the ST effect is activated. On the first LDT, we did not find evidence of a suppression effect since RT to threat words was not significantly slower than to neutral words. Conversely, a significant post-suppression rebound effect was found with a faster RT to threat words than to neutral words. This was particularly true for White participants under athletic-threat where both their RTs were significantly quicker to threat words compared to neutral words, and their RTs were significantly quicker to threat words compared to the RTs of Black participants in the same condition. Further support for a post-suppression rebound effect was found when directly comparing White athletic participants' quicker RTs to threat words on the post-suppression LDT to their RT to threat words on the suppression LDT.

This pattern is somewhat replicated in the Black IQ group. Although the results, were nonsignificant, the findings were in the predicted direction such that participants in the Black IQ group responded non-significantly faster to threat words than to neutral words, as expected. Further evidence for a post-suppression rebound effect is observed by the Black IQ group responding significantly quicker to threat words compared to White participants in the same condition. Finally, the Black IQ group responded non-significantly faster to threat words on the post-suppression LDT than on the suppression LDT. With adequate power, such findings would demonstrate strong evidence of a post-suppression rebound effect. Again, future studies need to

be conducted with a larger sample in order to better test the suppression and post-suppression rebound effect between Black and White participants under ST.

Limitations and Future Directions

Despite the study's interesting trends, the investigation was undermined by its smaller sample size and insufficient power. Given the pattern of results observed, it is warranted to invest in a larger dataset. Given the nonsignificant but consistent findings of an ST effect in both racial groups, it would be important to determine whether results are supportive of the well-known results in the field that Black participants experience ST on non-face-valid speed tasks. Additionally, greater power would be important to determine whether a double-dissociation of ST in Black-IQ and White-Athletic exists. A large sample would also be important to examining whether thought suppression could be ruled in or out as the mechanism if a ST effect were present.

Moreover, given the study's few statistical findings, moderating variables such as racial identification were not explored. Identification with one's race has been demonstrated to have both protective and deleterious moderating effects on ST salience and subsequent performance depending upon the stage one of one's own racial identity development (Davis et al., 2006). This line of research should be further explored on 'non-power' measures of intelligence and athleticism. It may be important in resolving the suggestion of some contradictory findings in current results.

References

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General, 130*(2), 224-237. doi:http://dx.doi.org/10.1037/0096-3445.130.2.224

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

Brown, R. P., & Pinel, E. C. (2003). Stigma on my mind: Individual differences in the experience of stereotype threat. *Journal of Experimental Social Psychology, 39*(6), 626-633. doi:http://dx.doi.org/10.1016/S0022-1031(03)00039-8

Browndyke, J.N, Albert, A.L, Malone, W., Schatz, P., Paul, R.H., Cohen, R., Cohen, R.A., Tucker, K.A., Gouvier, W.D. (2002). Computer-Related Anxiety: Examining the Impact of Technology-Specific Affect on the Performance of a Computerized Neuropsychological Assessment Measure. *Applied Neuropsychology, 9*(4), 210-218.

Cohen, G. L., & Garcia, J. (2005). "I am us": Negative stereotypes as collective threats. *Journal of Personality and Social Psychology, 89,* 566–582.

Davis, C., Aronson, J., & Salinas, M. (2006). Shades of threat: Racial identity as a moderator of stereotype threat. *Journal of Black Psychology*, *32*(4), 399–417. https://doi.org/10.1177/0095798406292464

Devine, A., Fawcett, K., Szűcs, D., & Dowker, A. (2012). Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety. *Behavioral and Brain Functions, 8*, 9. doi:http://dx.doi.org.ezproxy.lib.uwm.edu/10.1186/1744-9081-8-33

Everson, H. T., Smodlaka, I., & Tobias, S. (1994). Exploring the relationship of test anxiety and metacognition on reading test performance: A cognitive analysis. *Anxiety, Stress & Coping: An International Journal, 7*(1), 85-96. doi:http://dx.doi.org/10.1080/10615809408248395

Hopko, D. R., McNeil, D. W., Lejuez, C. W., Ashcraft, M. H., Eifert, G. H., & Riel, J. (2003). The effects of anxious responding on mental arithmetic and lexical decision task performance. *Journal of Anxiety Disorders, 17*(6), 647-665. doi:http://dx.doi.org/10.1016/S0887-6185(02)00240-2

Jensen, A. R. (2006). Clocking the mind: mental chronometry and individual differences. Amsterdam; Boston; London: Elsevier.

Johns, M., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resource depletion: Examining the influence of emotion regulation. *Journal of Experimental Psychology: General, 137*(4), 691-705. doi:http://dx.doi.org.ezproxy.lib.uwm.edu/10.1037/a0013834

Johns, M., Schmader, T., & Martens, A. (2005). Knowing is half the battle: Teaching stereotype threat as a means of improving women's math performance. *Psychological Science, 16*(3), 175-179. doi:http://dx.doi.org/10.1111/j.0956-7976.2005.00799.x'

Logel, C., Iserman, E. C., Davies, P. G., Quinn, D. M., & Spencer, S. J. (2009). The perils of double consciousness: The role of thought suppression in stereotype threat. *Journal of Experimental Social Psychology, 45*(2), 299-312. doi:http://dx.doi.org/10.1016/j.jesp.2008.07.016

Luce, R. D. (1986). Response times: Their role in inferring elementary mental organization. Oxford University Press

Medler, D.A., & Binder, J.R. (2005). MCWord: An on-line rrthographic database of the English language. http://www.neuro.mcw.edu/mcword/

Nguyen, H.-H. D., O'Neal, A., & Ryan, A. M. (2003). Relating test-taking attitudes and skills and stereotype threat effects to the racial gap in cognitive ability test performance. *Human Performance, 16,* 261–294.

Nisbett, Richard. (2011). The achievement gap: Past, present & future. Daedalus. 140. 90-100. 10.1162/DAED_a_00079

Oyserman, D., Kemmelmeier, M., Fryberg, S., Brosh, H., & Hart-Johnson, T. (2003). Racial-ethnic self-schemas. *Social Psychology Quarterly, 66*(4), 333-347. doi:http://dx.doi.org/10.2307/1519833

Pennington, Charlotte R., Heim, D., Levy, Andrew R., & Larkin, Derek T. (2016). Twenty years of stereotype threat research: A review of psychological mediators. *PLoS ONE, 11*(1), E0146487.

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

Sawyer, T. P., Jr., & Hollis-Sawyer, L. A. (2005). Predicting stereotype threat, test anxiety, and cognitive ability test performance: An examination of three models. *International Journal of Testing, 5,* 225–246.

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology, 38*(2), 194-201. doi:http://dx.doi.org.ezproxy.lib.uwm.edu/10 .1006/jesp.2001.1500

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85,* 440–452.

Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review, 115*(2), 336-356.

Schuster, C., Martiny, S. E., & Schmader, T. (2015). Distracted by the unthought - suppression and reappraisal of mind wandering under stereotype threat. *PLoS One, 10*(3) doi:http://dx.doi.org.ezproxy.lib.uwm.edu/10.1371/journal.pone.0122207

Smith, C. E., & Hopkins, R. (2004). Mitigating the impact of stereotypes on academic performance: The effects of cultural identity and attributions for success among African American college students. *Western Journal of Black Studies, 28,* 312–321.

Sommer, M., & Arendasy, M. E. (2014). Comparing different explanations of the effect of test anxiety on respondents' test scores. *Intelligence, 42*, 115-127. doi:http://dx.doi.org.ezproxy.lib.uwm.edu/10.1016/j.intell.2013.11.003

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

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:http://dx.doi.org/10.1037/0022-3514.69.5.797

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*(6), 1213-1227. doi:http://dx.doi.org/10.1037/0022-3514.77.6.1213

Wout, D., Danso, H., Jackson, J., & Spencer, S. (2008). The many faces of stereotype threat: Group- and self-threat. *Journal of Experimental Social Psychology, 44*(3), 792-799. doi:http://dx.doi.org/10.1016/j.jesp.2007.07.005

APPENDIX:

Orthographic Frequency and Word-String Length

Table A1. IQ Threat Words

String	Length	Frequency
complaining	11	10.7682
dishonest	9	5.88978
dumb	4	11.7201
ignorant	8	15.7656
illogical	9	2.20123
irrational	10	11.7201
lazy	4	12.7909
reckless	8	5.11638
threatening	11	18.5617
unmotivated	11	-
confused	8	29.5679
failure	7	68.6546
	M = 8.33	**M** = 17.52

Table A2. Athletic Threat Words

String	Length	Frequency
awkward	7	20.2275
blundering	10	1.54681
clumsy	6	12.2555
clutzy	6	-
indecisive	10	1.01138
rigid	5	24.3325
slow	4	81.1481
sluggish	8	3.62906
stiff	5	27.0692
stumbling	9	6.36572
uncoordinated	13	-
weak	4	47.9512
	M = 7.25	**M** = 22.55

Table A3 – Neutral Words

String	Length	Frequency
abbreviate	11	1.78
basement	9	13.15
daughter	8	103.46
federation	10	14.99
gleaming	8	11.01
kindly	6	23.74
ladder	6	13.68
libraries	9	9.28
onwards	7	9.28
plutonium	9	5.18
pneumonia	9	3.69
scarcely	8	32.72
	M = 8.33	**M** = 20.16

Table A4 – Non-words

String	Length
abrebviate	11
bsaement	9
dghautre	8
feerdation	10
gmelaing	8
kidnly	6
lededar	6
labrieries	9
ownards	7
plotunuim	9
pmeunonia	9
srcacely	8
	M = 8.33