Fall 2022

Exploring Social Identity Complexity and the Possible Impact on Stereotype Threat

Lindsay Quarles
Claremont Graduate University

Exploring Social Identity Complexity and the Possible Impact on Stereotype Threat

By
Lindsay Quarles

Claremont Graduate University

2022

APPROVAL OF THE DISSERTATION COMMITTEE

This dissertation has been duly read, reviewed, and critiqued by the Committee listed below, which hereby approves the manuscript of Lindsay D Quarles as fulfilling the scope and quality requirements for meriting the degree of Doctor of Philosophy in Psychology with a concentration in Applied Social Psychology.

Jason T. Siegel, Ph.D., Chair
Claremont Graduate University
Professor of Psychology

William Crano, Ph.D.
Claremont Graduate University
Oskamp Chair, Distinguished Professorship in Psychology

Allen M. Omoto, Ph.D.
Pitzer College
Vice President for Academic Affairs
Dean of Faculty

Marilynn Brewer, Ph.D.
Ohio State University
Emeritus Professor

Abstract

Stereotype threat is the fear of confirming as self-characteristic, a negative stereotype associated with a specific social identity that can hinder performance (Steele & Aronson, 1995). It is hypothesized that one way to protect against stereotype threat is to increase the availability of salient non-threatened identities (Croizet et al., 2001). Social identity complexity (SIC) is a theoretical construct that refers to the nature of the subjective representation of multiple ingroup identities (Brewer, 2008; Roccas & Brewer, 2002). The main hypothesis was that possessing a more complex social identity will provide protection against stereotype threat by providing increased availability of salient non-threatened identities. The findings of the three studies yielded contrasting results. With no consistent evidence, the conclusion is that there was too much variation introduced in the study due to the setting being largely removed from the stereotype relevant educational context. Additional analysis revealed that most African Americans indicated their race as the most or second most important aspect of their identity, indicating that racial identity is dominant and likely inextricably linked to their other identities. This finding suggests that the hypothesis of turning to another unthreatened identity is unlikely to ever buffer the effects of stereotype threat for African Americans when the identity under threat is their race. Ancillary analysis revealed self-affirmation to be effective in increasing test performance among African American participants, but this trend was the opposite for Latinx participants. Although these three studies could not definitively test the hypotheses, ancillary analyses revealed significant differences in identity attributable to race while controlling for survey location, suggesting that more research needs to be done to explore identity organization of racial minority members and how this impacts their experience of identity threat.

Keywords: stereotype threat, social identity complexity, self-affirmation

Dedication

I dedicate this dissertation to my daughter, the best part of me, Sanai. I rededicated myself to completing my degree because of you. I wanted you to know that you too can persevere through any challenge and accomplish the goals in your life. Never give up on your dreams and know that your parents will always be your safety net as you venture out on your own. As you watched me complete each step of this program, I was motivated by you. I wanted you to be proud of what your mom accomplished, as I was of my mother. Let my accomplishment of this next level propel you to even higher heights. I love you unconditionally and eternally.

Acknowledgements

I want to thank my mother, first and foremost for always supporting my dreams. She inspired me to pursue my dreams, showed me how to be ambitious, and provided a safety net for me to venture out into the world.

My supportive and loving husband, Kamal. Thank you for always having my back whether I wanted to quit my job to focus on completing my portfolio items or holding down the household while I escaped away to work on my dissertation. It was never easy, but you always believed in me and for that I will always be grateful.

My brother, Jay, who has always wowed me with his brilliant mind and passion for his work. You are an incredible teacher as you have even taught me a thing or two. You helped me to see that completion is possible, because if you could do it, I definitely could.

My grandmother, Grandy, who has always been there to share encouragement and a kick in the pants whenever I needed it.

Professor Siegel for taking on a student he barely knew and remaining a dedicated supporter of my work. Thank you for always providing a calming force when I am at my most stressed.

To my incredible friends, notably, Deryn, Valerie, and India who helped me to survive my first two years of graduate school with their friendship and laughter. To my friends who have provided support and love throughout this process, thank you Christina, Jamie, Tieranny, Justin, Tammie, and Ariana.

Contents

Exploring Social Identity Complexity and the Possible Impact on Stereotype Threat

Chapter 1:

Literature Review

Despite tremendous progress made to close the achievement gap between Black, Latinx and White students, significant disparities still exist. In 2018, on average White students earned 177 points higher on the SAT than Black students and 133 points higher than Latinx students (NCES, 2019). Studies into the Black/White achievement gap resulted in various explanations for its existence that vary along the nature to nurture spectrum. One widely accepted theory used to explain the achievement gap in academic performance between White and Black students is stereotype threat (Spencer et al., 1999; Steele, 2010; Steele & Aronson, 1995). "Stereotype threat is being at risk of confirming, as self-characteristic, a negative stereotype about one's group" (Steele & Aronson, 1995, p. 797). For African American students in a situation that requires them to demonstrate their academic abilities, such as a standardized exam, they fear confirming the stereotype that they are intellectually inferior, which can lead to subsequent underperformance (Inzlicht & Schmader, 2012).

To facilitate the development of more effective solutions for academic underperformance caused by stereotype threat, this dissertation explored the idea that one can protect against stereotype threat by increasing the availability of salient non-threatened identities (Croizet et al., 2001). The examination of this theory required exploration into the cognitive representations of one's social identities and their impact on performance under stereotype threat. To explore these ideas, it was important first to understand what stereotype threat is and how it works. Then explore a theory that explained individual conceptions of identity, which is hypothesized to impact the experience of stereotype threat. Last, to build on previous research that provided

evidence for successful interventions that alleviate the detrimental effects of stereotype threat, this dissertation examined if more complex conceptions of identity strengthen the buffering effect of self-affirmation when faced with stereotype threat.

Stereotype Threat Theory

Stereotype threat theory explained the persistent under performance of stereotyped group members as the result of subconscious automatic and controlled processes attributed to the threat of "confirming, as self-characteristic, a negative stereotype about one's group" (Steele & Aronson, 1995, p. 797). Dissemination of a negative stereotype about one's group produces a psychological threat, fear of being judged or mistreated according to the stereotype and such threat in an academic setting can hinder performance (Whitley & Kite, 2006). The pervasive detrimental performance effects of stereotype threat have been demonstrated in the underperformance of various stereotyped groups, such as African Americans on standardized exams (Steele & Aronson, 1995), women in math (Spencer et al., 1999), white men in athletics (Aronson et al., 1999), and Latinas' academic test performance (Gonzales et al., 2002). Specifically, African American students confront the stereotype that they are academically inadequate, as evidenced by historical works such as The Bell Curve that sought to support the prejudiced and inaccurate belief in Blacks' inferior intellectual ability (e.g., Herrnstein & Murray, 1994).

The groundbreaking research conducted by Steele and Aronson (1995) found evidence for stereotype threat in a study where they examined Black and White college students' performance on the verbal section of the Graduate Records Exam (GRE). Students were told either that the test was diagnostic (assesses intelligence) or non-diagnostic (assesses problem solving skills). Telling students that the test was diagnostic was intended to make their racial

group more salient, triggering the inferior intelligence stereotype (and associated poor performance) among Black students. In fact, Black students who were told the test was diagnostic performed significantly worse than their White counterparts. Those told the test assessed psychological factors involved in problem solving abilities performed just as well as their White counterparts.

The theory was then applied to a different stereotyped group, women in math, where Spencer et al. (1999) found evidence of detrimental performance effects associated with the negative stereotype of inferior mathematic ability among women. In this study, researchers tested the math performance of high performing math students, both male and female, and initially found that the women underperformed their male counterparts when the test was difficult, but not when it was less difficult. In this study participants were not intentionally provided a stereotype threat manipulation. They were given the normal instructions provided when completing the standardized exam. In the second study researchers tested for the possible stereotype threat present in the previous study by providing the difficult exam again, but instead of providing the normal testing instructions, half of the participants were told the test did not produce gender differences and half were told it did. Women who were told the test did not produce gender differences performed at the same level as their equally qualified male counterparts. However, women greatly underperformed the men when told the test yielded gender differences. These results demonstrated the significant negative effect stereotype threat has on the stereotyped group members.

Years later stereotype threat was first applied to Latinx students to test for the hypothesized *double-minority* effect of both race and gender stereotypes on mathematical and spatial test performance (Gonzales et al., 2002). In this study the researchers employed a 2

(diagnosticity: diagnostic, non-diagnostic) X 2 (ethnicity: Latino, White) X 2 (gender: male, female) between-individual factorial design to test for a possible "double-minority" effect of stereotype threat on Latina test performance. All participants were given the same paper-and-pencil test of numerical and spatial ability and those in the diagnostic condition were told that the study was investigating personal factors involved in performance, whereas those in the non-diagnostic condition were told the purpose was to understand the psychological factors involved in problem solving. Results revealed main effects of ethnicity and diagnosticity, in that White participants, and those in the non-diagnostic condition outperformed their counterparts. Analysis revealed there was no evidence of a possible "double-minority" effect in which women experienced greater ethnicity-based stereotype threat effects. Both Latino men and women were at risk of depressed performance due to the ethnicity-based stereotype threat effects in the diagnostic condition. In contrast, there was evidence of a "double-minority" effect in which Latinas experienced a greater performance decline in the diagnostic condition in comparison to Latinos. Latinas exhibited a large and statistically significant gender-based stereotype threat effect in comparison to White women who did not exhibit a statistically significant gender-based stereotype threat effect.

With the publishing of a multitude of stereotype threat research studies focused on both African American and Hispanic Americans, a meta-analysis was conducted to determine if there were statistically significant differences between the two populations or among the experimental methods used to create stereotype threat and their effects on outcomes (Nadler & Clark, 2011). The results of this meta-analysis found that African Americans and Hispanic Americans are affected similarly by stereotype threat as evidenced by the similarity in effect sizes between the two groups (d= .47 and d= .58, respectively). The meta-analysis also included a comparison of

implicit (simply making race salient) and explicit (direct threats such as mentioning stereotype-based expectations and hostile testing environments) methods used to create stereotype threat. Analyses revealed that there were no significant differences in effect sizes based on methodology of studies, providing evidence that stereotype threat can be activated effectively, using either implicit or explicit methods.

Underlying Processes

The underlying processes at work in stereotype threat are automatic and controlled effects stemming from the experience of being targeted by negative stereotypes (Schmader & Beilock, 2012). Each of the following processes provide explanations for stereotype-caused underperformance. These processes have since been theorized to represent connected mechanisms through which stereotype threat impacts performance. These processes are well established; however, they do not represent all the theories proposed to explain stereotype threat. It is important to note that being affected by stereotype threat does not require conscious awareness of the stereotype having been activated. There are several automatic responses that can hinder performance in the targeted domain regardless of awareness.

An overview of the stereotype threat process is that stereotyped group members under threat experience an automatic increase in stereotype activation, which leads to a cognitive imbalance in the relationships between the concepts of self, group, and ability in the domain that constitutes stereotype threat. The cognitive imbalance creates doubt that has been found to disrupt cognitive abilities (Schmader et al., 2009; Schmader et al., 2008). Researchers have also found evidence of a sense of uncertainty and increased vigilance toward negative cues when experiencing stereotype threat that result in hindered performance (Forbes et al., 2008).

Steele and Aronson (1995) found that Black college students who expected to complete an intelligence test demonstrated increased stereotype activation, in that they were more likely than their White peers to complete word fragments like R_C_ with the word RACE instead of other reasonable options such as RICE, ROCK, or RICH. These results revealed that describing the task as measuring one's intelligence brought the stereotype to mind. Having this negative stereotype activated, even subconsciously, is thought to lead to a cognitive imbalance in the relationships between the concepts of oneself, one's group, and ability in the domain, resulting in stereotype threat (Schmader et al., 2008). This imbalance is demonstrated in the proposed logical inconsistency between three relevant propositions; "I am a member of Group G, Group G is expected to do poorly at Domain D, but I do well at Domain D." Another possible inconsistency that constitutes stereotype threat is, "I am a member of Group G, Group G is expected to do poorly at Domain D, but I want to do well at Domain D." This inconsistency creates doubt that has been shown to disrupt cognitive abilities (Schmader et al., 2009). Schmader and Beilock (2012) explained that uncertainty is not the final result of stereotype threat, it is instead a phenomenological driver of additional processes aimed to resolve the inconsistency. As a result, one's attention becomes focused on cues that might provide evidence that one will have success at the task or not. However, since the goal is to avoid confirmation of the negative stereotype, attention is likely to be oversensitive in its detection of stereotype congruent cues. Therefore, any cue that would otherwise be innocuous, such as making a simple arithmetic error, can be over interpreted as a sign of failure. Forbes et al. (2008) found evidence for increased vigilance for negative cues when they tested minority college students who thought that their intelligence was being assessed. Using neurological measurements, research confirmed that participants had

increased vigilance to errors when they believed their intelligence was being assessed compared to when the task was described more neutrally.

In summary, situations of stereotype threat appear to cause an increase in automatic stereotype activation, resulting in inconsistent thoughts regarding oneself, one's group, and the targeted domain. This cognitive inconsistency triggers certain automatic effects such as a sense of uncertainty which results in increased vigilance toward negative cues. These effects are then met with more controlled processes aimed at managing one's behavior, thoughts, and emotions (Schmader & Beilock, 2012). Specifically, people experiencing stereotype threat have an increased motivation to do well, which results in increased effort at the task and decreased working memory.

An important detrimental controlled response elicited by stereotype threat is increased effort (Schmader & Beilock, 2012). Jamieson and Harkins (2007) added to the initial theory and explained that when people are threatened by how they may be evaluated, their increased motivation to perform well increases activation of the dominant response to the task. Unfortunately, the dominant response is detrimental to performance when the task is more cognitively challenging, but not when the task is simple. Ben-Zeev et al. (2005) found that women wrote their names faster repeatedly when they were told the subsequent math test had revealed gender differences in the past, compared to when they did not receive threatening information about the test. These results provide evidence that stereotype threat elicited a dominant response of name writing in an automatic way, which in the case of this simple task increased performance.

Jamieson and Harkins (2007) found evidence that women who were told the task was related to visuospatial and math ability were more likely to employ the dominant response,

which was contrary to the instructions, and then were more likely to make controlled attempts to correct the mistake. In study 4, participants completed two eye movement tasks, the antisaccade and prosaccade tasks, while indicating how often they saw a target letter appear on the screen that was either the same letter indicated by the researcher at the beginning of the study (0-back) or the same letter displayed on the screen two letters earlier in the sequence (2-back). Among participants in the 0-back condition, participants in the threat condition outperformed those in the no-threat condition, but among those in the 2-back condition, those in the threat condition performed more poorly than those in the control condition. These results revealed that adding a high-load concurrent task (2-back) that taxed working memory debilitated the performance of participants subject to threat on the antisaccade task, whereas stereotype threat alone facilitated performance when there was no or a minimal cognitive load (0-back).

Another well researched detrimental controlled response elicited by stereotype threat is decreased working memory. As stated previously, stereotype threat can increase one's motivation to perform well; however, it can decrease performance on tasks that require the mental manipulation of complex information (Schmader & Beilock, 2012). In the previously mentioned study conducted by Ben-Zeev and colleagues (2005) women performed well in the stereotype threat condition when the task was simple (quickly writing one's own name repeatedly). However, when faced with a more cognitively complex task (quickly writing one's own name backwards repeatedly) women in the stereotype threat condition performed worse than those not threatened. Researchers propose that performance under stereotype threat is impaired on more cognitively complex tasks because the threat strains working memory capacity (Beilock et al., 2003). Working memory can be thought of as a short-term memory system that is involved in the control, regulation, and active maintenance of a limited amount of information

with immediate relevance to the task at hand (Miyake & Shah, 1999). Working memory is also thought to allow one to focus attention on information relevant to the task, while inhibiting other irrelevant or distracting information (Engle, 2002). In a study conducted by Schmader and Johns (2003), women in the stereotype threat condition showed reduced working memory capacity and poorer math test performance relative to those in the control condition. In their analyses it was revealed that memory capacity mediated the link between stereotype threat and poorer math performance, providing evidence of a causal relationship.

Social Identity Complexity

Social identity complexity is a theoretical construct that refers to the nature of the subjective representation of multiple ingroup identities, the multiple group memberships in the overall social identity (Brewer, 2008; Roccas & Brewer, 2002). This construct was born from the commonly accepted idea that people belong to multiple groups, and therefore have some mechanism to conceptually organize their ingroup memberships (Roccas & Brewer, 2002). Social identity complexity describes the perceived overlap between different ingroups and these subjective representations range in complexity from one individual to another (Brewer, 2008; Roccas & Brewer, 2002; Schmid & Hewstone, 2011). This overview of social identity complexity will offer a detailed description of the construct, including its subcomponents and requirements.

Mechanisms of SIC

Social identity complexity varies on a continuum of complexity ranging from low to high (Brewer, 2008; Schmid & Hewstone, 2011). High complexity is described as having less perceived overlap between ingroups, whereas low complexity is having greater overlap. Take for example, an American woman who also is a Christian. If she perceived these groups as having

high overlap, meaning that the groups have the same members, she would have a relatively simple social identity. Her social identity is relatively simple because she does not perceive the existence of Americans who are not Christians, or Christians who are not American, and therefore does not include them in her ingroup (Roccas & Brewer, 2002; Schmid & Hewstone, 2011). In contrast, she would have a complex social identity if she realized that the groups do not always overlap, and included non-Christian Americans, and Christians of any nationality in her ingroup.

Social identity complexity can be divided into two distinct subcomponents, overlap complexity and similarity complexity (Schmid & Hewstone, 2011). Overlap complexity reflects the individual's perception of actual overlap in the numbers or proportions between different social categories (Roccas & Brewer, 2002; Schmid & Hewstone, 2011). Similarity complexity is the perceived similarity in definition, prototypically, or evaluative properties between categories (Schmid & Hewstone, 2011). For either, a high level of overlap reflects a relatively simple social identity (Roccas & Brewer, 2002). Therefore, an example of each is the idea that all African Americans are Democrats (demonstrating low overlap complexity) and then believing that the ideals of a Democrat are reflective of the ideals of African Americans (reflecting low similarity complexity).

There are two basic requirements needed for an individual to possess a relatively complex social identity. First, one needs to understand that there are multiple categories that make up a social identity. Second, one must understand that these different social categories do not, and most often cannot completely overlap in membership or similarity.

In summary SIC is a theoretical construct that refers to the nature of the subjective representation of multiple ingroup identities, and it varies on a continuum from simple to

complex (Brewer, 2008; Roccas & Brewer, 2002). The two subcomponents of social identity complexity are overlap, the proportion of overlap in membership, and similarity, the perceived similarity in definition, prototypically, or evaluative properties. To have a relatively complex social identity requires the acceptance of multiple categories that constitute one's social identity and the understanding that these categories do not, and most often cannot completely overlap in membership or similarity.

Proposed Influence of SIC on Stereotype Threat

Stereotype threat is the fear of confirming as self-characteristic a negative stereotype about one's group (Steele & Aronson, 1995). Therefore, given that SIC is the cognitive representation of one's social identities, it is proposed that an increased social identity complexity will buffer against the detrimental effects of stereotype threat by providing salient non-threatened identities that one can refer to when one identity is under threat. According to the self-concept literature, people possess a self-concept, a representation of the vast aspects of self-knowledge, and at any given time only some aspects of their self-concepts are accessible, constituting the working self-concept (Markus & Wurf, 1987). In this context stereotype threat can be conceptualized as targeted individuals' possession of both knowledge of the negative stereotype and the working self-concept associated with his or her stereotyped group that together can undermine performance. Based on this conceptualization one way to reduce the threat is to increase the accessibility of other, non-threatened aspects of one's self-concept (Croizet et al., 2001).

Several researchers have found evidence that support this hypothesis. Désert et al. (2001) found that when women experiencing stereotype threat were given the chance to describe themselves as a unique person, their performance was restored to the performance levels in the non-diagnostic condition. In a study conducted by Gresky et al. (2005) researchers found that

when faced with stereotype threat related to women's poor mathematic ability, women who highly identified with mathematics and drew self-concept maps that included many social roles and identities outperformed other women and performed equal or slightly better than their male counterparts. Carmichael (2011) revealed a trend, though not significant, that suggested that women with high social identity complexity outperformed women with low social identity complexity when faced with stereotype threat associated with women in mathematics. These results provide support for the hypothesis that social identity complexity may improve performance by reducing the salience or importance of the stereotyped identity and raising the salience of a non-threatened identity.

CHAPTER 2:

The Current Studies

Based on reasoning from the social identity complexity literature, the current studies examined the impact, if any, varying levels of social identity complexity had on the experience of stereotype threat. The idea at the center of the current studies is that possessing a more complex social identity allows one to avoid the detrimental performance outcomes for stereotyped group members by increasing one's availability of non-threatened social identities, thus buffering against the effects of stereotype threat. Additionally, this dissertation research assessed the possible additive effect of increased social identity complexity in conjunction with an effective stereotype threat intervention, self-affirmation. These three studies examined three hypotheses:

1) People with higher social identity complexity will outperform those with lower social identity complexity when faced with stereotype threat,

2) Social identity complexity can be manipulated, and when increased before experiencing stereotype threat can have a buffering effect, and

3) Bolstered social identity complexity will increase the effectiveness of self-affirmation in improving performance under stereotype threat.

In the current studies the impact of social identity complexity was assessed in two of the most studied stereotyped groups, Black and Latinx students. Since Steele and Aronson's (1995) research first revealed evidence of the detrimental effects of stereotype threat, numerous studies have supported their original hypotheses. Most studies conducted with Black students were conducted among college students, and the aim of this research is to explore a wider spectrum of students, including those who did not enroll in a four-year college or university.

This research was composed of three studies. The first tested the first hypothesis, that people with higher social identity complexity will outperform those with lower social identity complexity when faced with stereotype threat. To do this, in study 1 participants were randomly assigned to either the stereotype threat or control condition and completed an academic test. After the test, students completed a social identity complexity measure (Brewer & Pierce, 2005). This allowed for a comparison of performance between students with higher and lower social identity complexity under threat. If hypothesis 1 was confirmed, it would provide evidence that variation in social identity complexity has significant implications for performance under threat.

Study 2 tested the second hypothesis, that social identity complexity can be manipulated, and when increased before experiencing stereotype threat, it can have a buffering effect. To test this hypothesis, students were randomly assigned to complete an exercise designed to increase their social identity complexity (Gresky et al., 2005) or were assigned to the control condition, then students were randomly assigned to either the stereotype threat or control condition, completed an academic test, and then the social identity complexity measure. If this hypothesis was supported it would provide strong evidence that social identity complexity can be manipulated, supporting the original theoretical reasoning proposed by Brewer and Roccas (2002). Secondly, these results could also demonstrate an effective strategy to buffer stigmatized group members against experiencing the detrimental effects of stereotype threat.

Study 3 explored the additive effect of bolstering social identity complexity in conjunction with a commonly used stereotype threat intervention, self-affirmation. Self-affirmation has been found to buffer against the detrimental performance outcomes associated with stereotype threat, by asking participants to reflect on important aspects of their lives or personal values that are often different from the threatening domain prior to completing a

stereotype threatened task (e.g., Cohen et al., 2006; Martens et al., 2006; Shapiro et al., 2013). Researchers found that Black seventh grade students from middle- to lower middle-class families who completed a self-affirmation exercise had higher grades at the end of the term than those who completed a neutral exercise (Cohen et al., 2006). Women who were faced with stereotype threat and affirmed a value attribute before taking a diagnostic test performed similarly to those in the control condition and outperformed those in the stereotype threat condition who did not affirm a value attribute (Martens et al., 2006). Therefore, in study 3, self-affirmation was included to determine if social identity complexity can bolster performance when increased in conjunction with receiving the intervention, while faced with stereotype threat.

In study 3 students were randomly assigned to receive the social identity complexity exercise or the control condition, next students were randomly assigned to either the stereotype threat or control condition, and then randomly assigned to receive the self-affirmation intervention or the control condition. Lastly, the students completed the academic test and social identity complexity measure. If the hypothesis were supported, it would provide evidence that social identity complexity not only buffers against stereotype threat, but it can also amplify the positive effects of an existing intervention.

CHAPTER 3:

Study 1

Hypothesis

Study 1 tested the first hypothesis, that people with higher social identity complexity would outperform those with lower social identity complexity when faced with stereotype threat. If this hypothesis were confirmed, it would provide the foundational evidence that variation in social identity complexity has significant implications for academic performance under threat.

Methods

Participants

Based on power analysis using G*Power (Faul et al., 2007), the minimum total sample size of approximately 98 (approximately 49 in each group) was necessary to have .8 power to detect a moderate effect (i.e., f^2= .15). The moderate effect size, d= .8, was determined based on the meta-analyses that determined that generally stereotype threat has a moderate effect on intellectual performance among African American and Latinx participants (Nadler & Clark, 2011). Additionally, the effect size observed in the study where participants' social identity complexity was manipulated to buffer math test performance under threat among women college students also was moderate, d = .7 (Gresky et al., 2005).

Participants included were 213 adults (age 18 and over) recruited through a variety of mechanisms. This study was designed for college students, however, with the impact of COVID-19 making most college settings remote, the study recruitment was adjusted to include the online survey recruitment platform, Prolific. Most participants were recruited from a large mid-western community college (130), with one participant recruited through the Black Student Association at Claremont Consortium, and 82 recruited through Prolific. Participants recruited through the

colleges were offered entry in a raffle as an incentive to participate. Participants recruited through the colleges each had a one in twenty chance to win a raffle prize of a $10 gift card to local a movie theatre, Chipotle, or Target. Participants recruited through Prolific were offered $3.75 and told the survey would take approximately 25 minutes of their time. Participants recruited through Prolific were prescreened to allow the survey to be viewed only by respondents who were between the ages of 18-40, Nationality: United States, Ethnicity: Black/African American, Latino/Hispanic, and having at least a high school diploma.

Of the 213 respondents who completed the survey, 97 were included in the final analyses based on meeting the racial inclusion criteria of identifying as either African American/Black (50.5%) or Hispanic/Latinx (49.5%) and identifying as either male or female. The demographics of the 116 respondents removed from analyses were as follows: Asian/Pacific Islander (1), White/Caucasian (95), Other (3), Multi-racial (13), and Native American (1). Two Black/African American and one Latinx participants were removed from analysis because they did not identify as either male or female. Of the 97 participants included in the analyses, 58.8% identified as women, 41.2% identified as men. The majority of participants included in the analyses (52.6%) were enrolled in a post-secondary institution when completing this survey. The sample used for analysis was comprised of participants recruited through Prolific (80.4%) and from the large midwestern community college (19.6%). The age range of participants was not captured among the participants recruited through the colleges, so their age ranges are unknown (19.6%). The age ranges of the participants recruited through Prolific are as follows: 18-24 (35.1%), 25-32 (36.1%), 33-40 (9.3%).

Design and Procedure

The study employed a mixed experimental design with one manipulated variable, threat condition (stereotype threat/diagnostic or control/non-diagnostic). The survey was completed online for all participants using Qualtrics as the survey platform. Upon entering the survey, the first page displayed the informed consent form and participants had to check a box indicating their consent to continue to the rest of the survey. After consenting participants were then randomly assigned to a threat condition that was delivered through the test instructions. This manipulation was adapted from Steele and Aronson (1995) and focused on the test as a measure of one's own abilities, indicating a self-as-target stereotype threat (Shapiro & Neuberg, 2007). The test instructions for those assigned to the stereotype threat condition read as:

> Today you will have 15 minutes to complete a test of your mathematical and verbal abilities. This is a difficult test and you should not expect to get many items correct. The difficulty of this test provides a genuine measure of your mathematical and verbal abilities so that I might better understand your strengths and weaknesses. Please try your best on this exam in order to accurately measure your mathematical and verbal abilities. After completing the test you will be provided with feedback about your performance in comparison to other people.

The instructions for those assigned to the control condition read as follows:

> Today you will have 15 minutes to complete an assignment used to better understand problem-solving. This is a difficult assignment and you should not expect to get many items correct. Please do your best even though this is not a test of your ability. Again, this is not a test of your ability in any way. This information will help me in my analysis of the problem-solving process.

After reading the instructions all participants completed a multiple choice 10-question test comprised of five math and five verbal questions from a practice SAT exam. Participants who were recruited through the colleges completed a manipulation check that assessed their belief about the task; however, those who were recruited via Prolific did not as it was unintentionally left off. Once completed all participants completed a social identity complexity (SIC) measure (Appendix A). The social identity complexity measure was provided after the stereotype threat manipulation and test to avoid any risk of impeding the stereotype threat effect. Then participants finished with responses to a demographics survey and were debriefed. After the debriefing participants in the colleges sample were offered a link to another survey form where they could provide their email address and prize preference for the raffle. Having the raffle survey separate from the study survey allowed for study survey responses to be unmatchable to the personally identifiable information needed for the raffle.

Measures

Academic test. The 10-item test consisted of 5 math and 5 verbal problems from the publicly available practice SAT. Reliability assessment revealed internal test-retest Cronbach alphas for both the math and verbal domains; $r(97)= .54$ and $r(97)=.54$, respectively. Collectively the 10-item test had a Cronbach's alpha of $\alpha = .58$. Although the coefficient was below the acceptable level of .7, this test was not intended to serve as a measure of intellectual ability, but instead a challenging task used to trigger and measure the stereotype effect.

Manipulation check. Drawn from Steele and Aronson (1995), the manipulation check asked participants to select the purpose of the experiment:

The purpose of this experiment was to:

(a) Provide a genuine test of my mathematical and verbal abilities; or

(b) Provide a difficult assignment to examine factors involved in problem solving.

Social identity complexity measure. Adapted from Miller et al. (2009), in this measure of overlap complexity participants are probed for domains of social identities, including political organizations, gender identity, recreational and sports groups, religious affiliation, and ethnic/national identities, then asked to choose four that are most important to them. Next, each of the four groups selected are paired and the participant is asked to estimate the overlap in membership between the groups as they perceive them (i.e., "How many (Group A) are also (Group B)?") Participants responded using a scale of 0 (none are) to 10 (all are) for all pairings, in both directions, for a total of 12 overlap ratings. An overlap complexity score is computed as the mean of these 12 ratings, with high overlap scores indicating low social identity complexity. Please see the social identity complexity measure used in the study in Appendix A.

Demographics and debriefing. The demographics section for the college recruited participants included information regarding participants' race, gender, grade level, academic aspirations and whether they completed the study in the classroom or elsewhere. The demographics section for the Prolific recruited participants included race, gender, US residency status, postsecondary enrollment, grade level, highest educational level attainment, and age range. Once completed all participants were debriefed and thanked for their participation.

Results

Data Management

Participants assigned to the threat/diagnostic condition were coded with a threat condition score of "2" and participants assigned to the control/non-diagnostic condition were coded with a threat condition score of "1." Lower scores on the SIC measure indicate high social identity complexity and higher scores indicate low social identity complexity. Black participants were

coded with a "1" for Race and Latinx participants were coded with a "2." Men were coded with a "2" for Gender and Women were coded with a "1." The two survey locations were dummy coded as Prolific to indicate participants recruited through Prolific with those not included indicating recruitment through the community college.

Preliminary Analyses

Table 1

Range, Mean Scores, Standard Deviations, and Correlations for Sociodemographic Variables, Independent, and Dependent Variables

		Range	*M*	*SD*	1	2	3	4	5
1	Test Score	1-10	6.18	2.15					
2	Threat Condition	1, 2	1.51	0.50	-0.04				
3	SIC Score	2-10	5.31	1.59	-0.20*	-0.02			
4	Race	1, 2	1.49	0.50	0.22*	0.03	-0.01		
5	Gender	1, 2	1.41	0.49	0.06	-0.09	-0.02	-0.16	
6	Prolific	0, 1	0.80	0.40	0.22**	-0.02	-0.15	-0.39**	0.10

Notes. * Correlation is significant at the 0.05 level (1-tailed)

** Correlation is significant at the 0.01 level (1-tailed)

Correlational analyses of the independent, dependent, and demographic variables revealed that the survey location (college or Prolific) is correlated with test scores, $r = .22$, $p = .01$ and participants race, $r = -.39$, $p < .01$. ANOVA analysis was conducted with race, survey location and an interaction term to examine the relationships. This analysis revealed a significant main effect of survey location, $F(1,93) = 4.12$, $p = .05$, where participants recruited via Prolific

(*M* = 6.41, *SD* = 2.16) outperformed the sample recruited from the community colleges (*M* = 5.21, *SD* = 1.87). A review of the proportion of participants by race and survey location reveal that the college sample was disproportionately African American, therefore the subsequent Prolific sample was intentionally disproportionately Latinx to balance the total sample (see Table 2). Lastly, the correlational analysis reveal that SIC is related to test performance, *r* = -.20, *p* = .03, revealing that those with more complex social identities (coded as having lower SIC scores) demonstrated higher test scores. This is generally consistent with the hypothesis that those with higher social identity complexity would perform better on the test, but the hypothesis is specific to those faced with stereotype threat and this correlation does not include threat.

Table 2
Sample Race by Survey Location

Race	College Sample	Prolific Sample	Total
Black/African American	17	32	49
Latinx/Hispanic	2	46	48

Main Analysis

The hypothesis for study 1 was that people with higher social identity complexity will outperform those with lower social identity complexity when faced with stereotype threat. To test this hypothesis, a hierarchical multiple regression was conducted with 48 participants randomly assigned to the control/non-diagnostic condition and 49 participants randomly assigned to the threat/diagnostic condition. In the first step of the regression were the demographic variables: gender, race, and survey location, then the SIC measure on the second step, on the third step the threat manipulation, and the last step included an interaction variable for threat manipulation and the SIC measure. Due to the meta-analysis conducted by Nadler & Clark (2011) that revealed both African American and Hispanic Americans people are affected

similarly by stereotype threat race was not included as a predictor of test performance and instead treated as a control variable.

The results of the regression analysis revealed the final model to explain 14.6% of variability of test performance, $F(6, 90) = 2.57$, $p = .02$ (Table 3). Step 2 of the regression which included the demographic variables and SIC measure explained 10.4% of variability of test performance, $F(4, 92) = 2.67$, $p = .04$. The SIC measure contributed little to the model, providing no statistically significant evidence that it alone effects test performance, $\Delta R^2 = .03$, $\Delta F(1,92) = 3.01$, $p = .09$. A closer look at the stereotype threat manipulation effect, entered on the third step, showed that the threat condition did not explain any of the variance in test scores, $\Delta R^2 = .00$, $\Delta F(1,91) = 0.19$, $p = .66$.

Table 3

Hierarchical Multiple Regression Results Predicting Test Performance from Threat Condition and SIC

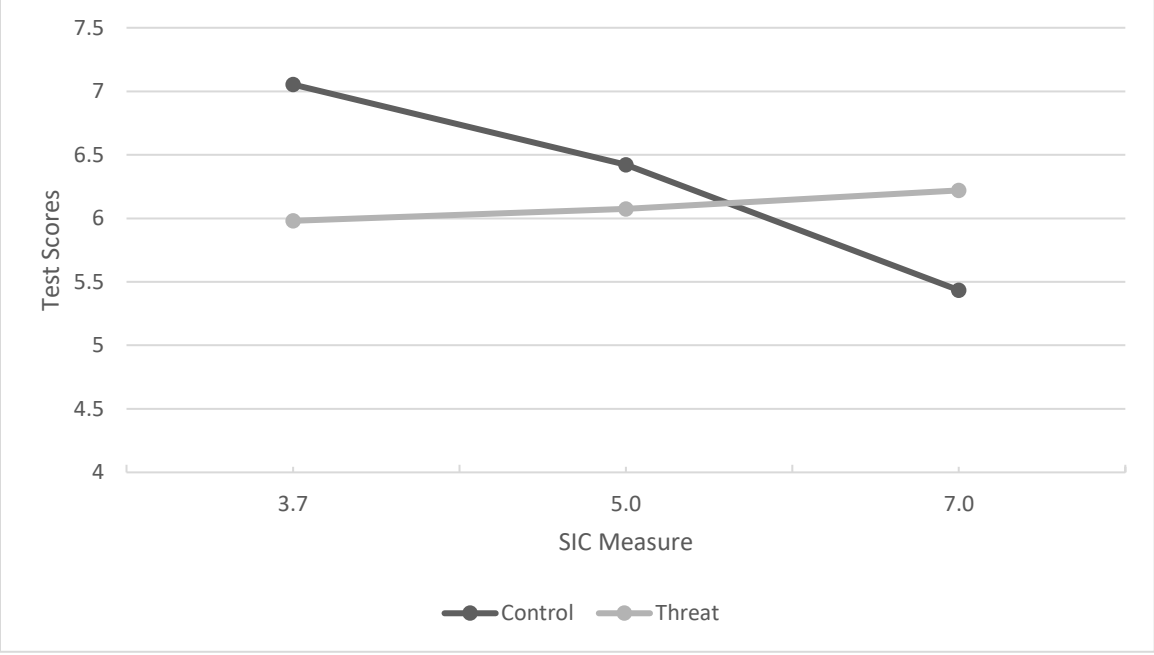
Predictor	Step 1			Step 2			Step 3			Step 4		
	B	SE_B	β	*B*	SE_B	β	*B*	SE_B	β	*B*	SE_B	β
Race	.73	.47	.17	.77	.47	.18	.78	.47	.18	.86	.47	.20
Gender	.31	.45	.07	.31	.44	.07	.30	.45	.07	.43	.44	.10
SurvType	.81	.59	.15	.64	.59	.12	.64	.60	.12	.85	.59	.16
SIC				-.24	.14	-.17	-.24	.14	-.17	-1.06	.42	-.79*
Threat							-.19	.43	-.04	-3.19	1.51	-.74*
ThreatXSIC										.57	.28	.95*
R^2		.08			.10			.10			.15	
F		2.51			2.67			2.16			2.57	
p		.06			.04			.07			.02	

Notes. * $p < .05$

The final model revealed a significant interaction effect in that test scores among participants in the threat/diagnostic condition were largely unrelated to SIC scores ($\Delta R^2 = .04$, $\Delta F(1,90) = 4.27$, $p = .04$); whereas participants in the control condition who possessed more complex social identities displayed higher test scores ($B = 0.57$, $SE = 0.28$, $p = .04$). These two findings do not support the hypothesis that possessing a more complex social identity when faced with stereotype threat would buffer against the detrimental effects of stereotype threat. Not only did SIC not have the expected relationship with test scores among those in the threat/diagnostic condition, but there was not expected to be a relationship between SIC and test performance in the control condition. The interaction is displayed in Figure 1. This relationship was expected to exist among those in the threat condition and not the control condition.

Figure 1

SIC and Threat Manipulation Interaction on Test Scores.



Discussion

The results of study 1 do not support the hypothesis that people with higher social identity complexity will outperform those with lower social identity complexity when faced with stereotype threat, as evidenced by the significant interaction effect that revealed participants faced with stereotype threat were largely unaffected by SIC. This result does not support the idea that possessing a more complex social identity would help one to buffer the detrimental effects of an identity threat by turning to another non-threatened identity. However, the results of this study did reveal a positive relationship between SIC (higher complexity is demonstrated by lower SIC scores) and test performance among those in the control condition. Although the hypothesis focused on the impact of SIC on those in the threat/diagnostic condition, it is somewhat consistent with the hypothesis that those with higher SIC would demonstrate higher test scores because there was a positive relationship between SIC and test scores, even though it was in the control condition.

The stereotype threat manipulation only had an effect once the interaction term was included in the regression. This was an unexpected outcome as the stereotype threat manipulation was expected to have a more significant impact on test performance based on previous findings in literature that found on average a moderate effect size (Nadler & Clark, 2011). In this study there was no significant main effect of stereotype threat, and it is unclear if that is simply due to the introduction of social identity complexity or if it is due to a failure of the threat manipulation to elicit the full strength of its intended effect.

Study 1 was designed to test for a relationship between social identity complexity and test performance under stereotype threat among Black and Latinx people. The theory being tested is that people who have more complex social identities would be able to turn to another identity

when one is under threat, thereby protecting themselves from the identity threat. This first study was a simple experimental design based on the research from Roccas and Brewer (2002) that people naturally vary on a continuum from simple to complex social identities, and analysis was designed to simply measure the relationship, establishing a foundation for this new theory. However, the results do not support this hypothesis, but instead suggest that there is no relationship between SIC and stereotype threat. As the remaining two studies build on the hypothesis in this study it is possible that the subsequent studies will not have evidence to support their hypotheses. The subsequent studies will replicate the procedures of study 1, allowing for a test of replication to determine if the contrasting findings in this study are the result of Type 1 error or evidence of a relationship between social identity complexity and academic performance not previously expected.

Limitations

There were three significant limitations of this study. The first was the unintended omission of a manipulation check for the stereotype threat effect. The omission of this manipulation check eliminated an opportunity to further explore the stereotype threat manipulation and understand why it didn't have the expected effect.

The second limitation was the significant differences in test performance associated with survey location. The addition of Prolific participants was meant to provide a sufficient sample size to test the hypothesis, but it resulted in confounding race with survey location as the participants were not evenly distributed across locations.

The third significant limitation of this study was in conducting it virtually and removing it from the relevant academic setting. It is possible that the stereotype threat manipulation did not

work as expected because participants were removed from the relevant setting and completing a task that had no significance to the participants' lives beyond the study.

CHAPTER 4

Study 2

Hypothesis

Study 2 tested the second hypothesis, that social identity complexity can be manipulated, and when increased before experiencing stereotype threat, can have a buffering effect. If this hypothesis is confirmed it would provide some evidence that there is a relationship between social identity complexity and test performance when under stereotype threat. Additionally, if this hypothesis is confirmed, it would provide evidence that social identity complexity can be changed and is not fixed. If this worked it could provide a potentially effective and simple intervention to help Black and Latinx students improve their academic performance in school, thus increasing the opportunities available to them in the future.

Methods

Participants

Based on power analysis using G*Power (Faul et al., 2007), a sample size of approximately 127 (approximately 32 in each group) will be necessary to obtain .8 power to detect a moderate effect (i.e., f^2= .15). The moderate effect size, d= .8, was determined using the same reasoning as detailed in study 1, as this study was designed to be a continuation of the previous research.

Participants included in study 2 were 364 adults (aged 18 and over) recruited through a variety of mechanisms. As in study 1 with the impact of COVID-19 making most college settings remote, the study recruitment was adjusted to include the online survey recruitment platform, Prolific. Again, most participants were recruited from a large mid-western community college (287), with 250 completing the in the classroom and 37 completing it outside the classroom. Fifteen participants were recruited through a large city community college in the

northeast, with 8 completing the survey in the classroom and 7 completing it outside the classroom. Lastly, 62 were recruited through Prolific. As in study 1, participants recruited through the colleges were offered entry in a raffle as an incentive to participate. Participants recruited through Prolific were offered $5.25 and told the survey would take approximately 35 minutes of their time. Participants recruited through Prolific were prescreened to allow the survey to only be viewed by respondents who were between the ages of 18-40, Nationality: United States, Ethnicity: Black/African American, Latino/Hispanic, having at least a high school diploma, and did not previously participate in study 1.

Of the 364 respondents who completed the survey, 140 were included in the final analyses based on meeting the racial inclusion criteria of identifying as either African American/Black (61.4%) or Hispanic/Latinx (38.6%) and identifying as either male or female. The demographics of the 224 respondents removed from analyses were as follows: Asian/Pacific Islander (9), White/Caucasian (181), Other (16), and Multi-racial (16). Of the 140 participants included in the analyses, 57.9% identified as women and 41.2% identified as men. The majority of participants (77.9%) were enrolled in a post-secondary institution when completing this survey. As the original data collection plan was to recruit college students, the Prolific recruitment was done subsequently to meet the sample minimum, resulting in the sample composition of 61.4% recruited from the two community colleges and the remaining 38.6% recruited from Prolific. The age range of participants was not captured among the participants recruited through the colleges, so their age ranges are unknown. The age ranges of the participants recruited through Prolific are as follows: 18-24 (16.2%), 25-32 (19.0%), 33-40 (3.5%).

Design and Procedure

The study employed a 2 (threat condition: stereotype threat/diagnostic or control/non-diagnostic) X 2 (SIC manipulation: self-concept map or toy concept map) experimental design. As in study 1, the survey was completed online for all participants using Qualtrics as the survey platform. Upon entering the survey, the first page displayed the informed consent form and participants had to check a box indicating their consent to continue to the rest of the survey.

After consenting, participants were then randomly assigned to a SIC manipulation condition, adapted from Gresky et al. (2005). In the self-concept map condition participants were asked to create a self-concept map where they wrote the word "Me" in the center and then were asked to draw lines and nodes connecting to other identities (e.g., school, family, interests). See Appendix B for an illustration of the SIC manipulation for participants in the experimental condition asked to create self-concept maps. This task was designed to increase social identity complexity by allowing participants to think about a variety of components of their social identity. In the control condition participants were provided similar instructions and asked to draw a concept map of the word "toys," designed to have no effect on their social identity complexity See Appendix C for an illustration of the SIC manipulation for participants in the control condition asked to create concept maps. In both the self-concept map and control condition participants were asked to use their own writing utensil and sheet of paper to complete the task.

After participants completed their concept maps, participants were instructed to email or text a picture of the concept map to the researcher before proceeding to the next section of the survey. Once completed, participants were then randomly assigned to the stereotype threat/diagnostic condition or control/non-diagnostic condition as in study 1. After receiving the

threat condition, participants completed the verbal and math test, manipulation check and the social identity complexity measure, also as in study 1. After completing the social identity complexity measure participants completed a Domain Identification Measure (DIM) also adapted from Gresky et al., included as Gresky and colleagues found the effect of drawing self-concept maps was more effective among those that identified with academics. Then participants finished with responses to a demographics survey and were debriefed, and those in the colleges sample were offered entry into the raffle, just as in study 1.

Measures

Academic test. Replicated from study 1, the 10-item test consisted of five math and five verbal problems from the publicly available practice SAT. Reliability assessment revealed internal test-retest Cronbach alphas for both the math and verbal domains; $r(140)=.18$ and $r(140)=.61$, respectively. Collectively the 10-item test had a Cronbach's alpha of $\alpha = .56$. Although the coefficient is below the acceptable level of .7, this test was not intended to serve as a measure of intellectual ability, but instead a challenging task used to trigger and measure the stereotype effect.

Manipulation check. As in study 1, drawn from Steele and Aronson (1995), the manipulation check asked participants to select the purpose of the experiment:

The purpose of this experiment was to:

(c) Provide a genuine test of my mathematical and verbal abilities; or

(d) Provide a difficult assignment to examine factors involved in problem solving.

Social identity complexity measure. As in study 1, the social identity complexity measure was adapted from Miller et al. (2009), participants are probed for domains of social identities, including political organizations, gender identity, recreational and sports groups,

religious affiliation, and ethnic/national identities, then asked to choose four that are most important to them.

Domain Identification Measure (DIM). In Gresky and colleagues' (2005) study the exercise used to increase social identity complexity was most effective among women who were highly identified with math. Therefore, in the case that the exercise is found to be ineffective across the general sample, this domain identification measure may provide an explanation. The domain identification measure drawn from Smith and White (2001) assesses identification within the mathematics and English domains. The 16-item questionnaire was estimated to take approximately five-ten minutes to complete and contains items such as:

"I get good grades in Math" (on a scale from 1 = *strongly disagree* to 5 = *strongly agree*)

"How much do you enjoy English-related subjects?" (On a scale from 1 = *not at all* to 5 =

very much)

Reliability assessment revealed internal test-retest Cronbach alphas for both the math and English domains; $r(140) = .92$ and $r(140) = .84$, respectively. Collectively the 16-item test as a measure of identification with academics and has a Cronbach's alpha of $\alpha = .83$.

Demographics and debriefing. The demographics section for both the college and Prolific recruited participants was the same as in study 1. The demographics section for the college recruited participants included information regarding participants' race, gender, grade level, academic aspirations and whether they completed the study in the classroom or elsewhere. The demographics section for the Prolific recruited participants included race, gender, US residency status, postsecondary enrollment, grade level, highest educational level attainment, and age range. Once completed all participants were debriefed and thanked for their participation.

Results

Preliminary Analyses

Correlational analyses of the independent, dependent, and demographic variables revealed that two of the survey locations (community college 1 in-person and Prolific) are each correlated with participants' race, $r = -0.37$, $p < .001$ and $r = 0.37$, $p < .001$, respectively (see Appendix D). A review of frequency data reveals that the sample recruited via community colleges were disproportionately African American. Again, the subsequent Prolific sample was intentionally disproportionately Latinx to balance the total sample (Table 4). Correlational analysis revealed that the race of the participant is related to their performance on the test, $r = 0.30$, $p < .001$. ANOVA analysis was conducted with race, survey location and an interaction term to examine the relationships. This analysis revealed a significant main effect of survey location, $F(4,130) = 3.13$, $p = .02$, and no main effect of race or interaction effect. This finding demonstrates that race and survey location are confounded and that the differences in test performance associated with race are likely attributable to variances associated with the different survey locations.

Table 4

Race by Survey Location

Race	College Sample	Prolific Sample	Total
Black/African American	*65*	*21*	*86*
Latinx/Hispanic	*21*	*33*	*54*

Correlational analysis showed that participants' race was related to the threat manipulation they received, in that African American participants were more likely to be assigned to the stereotype threat/ diagnostic condition, $r = -0.18$, $p = .03$ (see Appendix E).

Main Analysis

The hypothesis for study 2 was that social identity complexity can be manipulated, and when increased before experiencing stereotype threat can have a buffering effect, thus protecting impacted people from experiencing the detrimental effects on academic performance. To test this hypothesis, a hierarchical multiple regression was conducted. In the first step of the regression were the demographic variables: gender, race, and survey location, then the Domain Identification Measure (DIM) on the second step. On the third step was the threat manipulation, followed by the SIC measure on step four. Then on the fifth step of the hierarchical regression the SIC manipulation was introduced. The last two steps included four interaction variables: (1) threat manipulation and the SIC measure, (2) threat manipulation and SIC manipulation, (3) SIC measure and SIC manipulation, and (4) DIM and threat manipulation., with a three-way interaction between SIC measure, threat manipulation, and SIC manipulation as the final step.

Although the results of the regression analysis revealed the final model to explain 25.9% of variability of test performance, $F(15, 117) = 2.73$, $p = .001$, though the first step of the regression, which only included the demographic variables was the only one to significantly improve the model prediction ($\Delta R^2 = .20$, $\Delta F\ (6,126) = 5.36$, $p < .001$). These results did not support the hypothesis that identity complexity can be manipulated, and when increased before experiencing stereotype threat, it has a buffering effect. There was no effect of the SIC measure ($B = -.04$, $SE = .10$, $p = .71$) or the SIC manipulation ($B = -.12$, $SE = .35$, $p = .73$) when entered in the fourth and fifth steps of the regression respectively (Table 5). This result showed that neither the self-concept maps nor participants' social identity complexity were significant predictors of participants' test performance. Utilizing ANCOVA for a closer look at the impact of the SIC manipulation, while controlling for race, gender and survey location, revealed that it

had no significant impact on participants' level of social identity complexity as indicated by the SIC measure, $F(1,135) = 1.41$, $p = .24$. These results show that drawing a self-concept map had no effect on social identity complexity, thus the manipulation failed to elicit the expected effect. Although there was a main effect of race, it is important to note that this variable is confounded with survey location which also revealed a significant main effect.

Table 5

Hierarchical Multiple Regression Results Predicting Test Performance from Threat Condition and SIC Manipulation

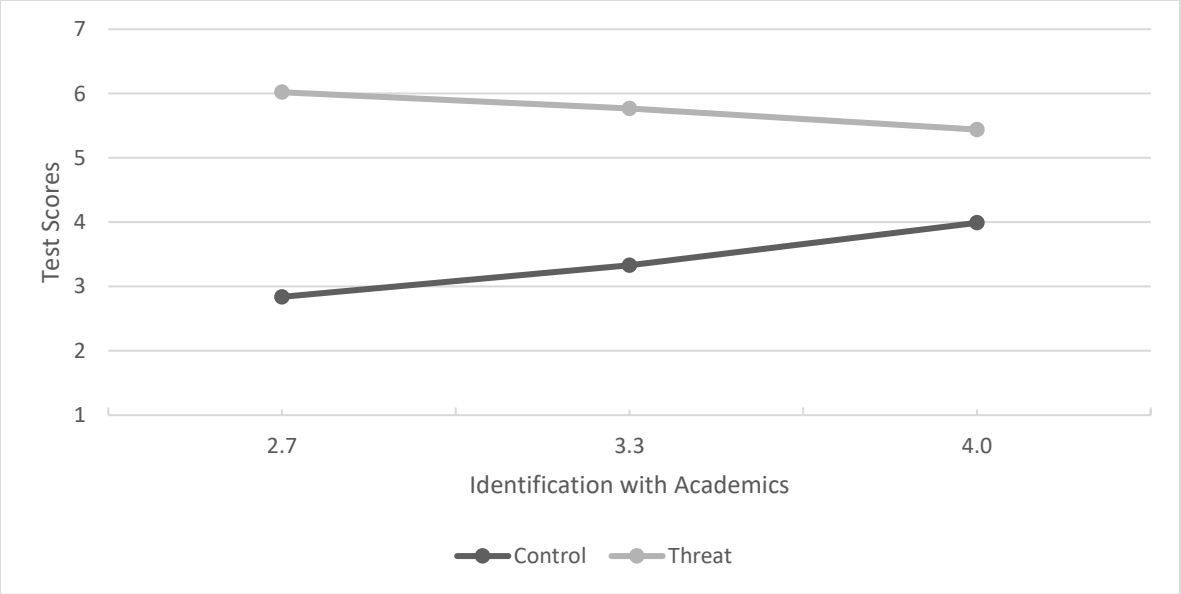
Predictor	Step 1			Step 2			Step 3			Step 4			Step 5			Step 6			Step 7		
	B	SE_B	B	*B*	SE_B	B	*B*	SE_B	β	*B*	SE_B	β	*B*	SE_B	β	*B*	SE	β	*B*	SE_B	β
Race	1.12	.37	.26**	1.13	.37	.26**	1.16	.38	.27**	1.16	.38	.27**	1.16	.38	.27**	1.24	.38	.29**	1.24	.38	.29**
Gender	.31	.34	.07	.26	.35	.06	.26	.35	.06	.25	.36	.06	.27	.36	.06	.36	.36	.08	.37	.36	.09
CC1 In-person	-1.41	1.01	-.12	-1.45	1.01	-.12	-1.45	1.02	-.12	-1.41	1.03	-.12	-1.40	1.03	-.11	-1.45	1.06	-.12	-1.50	1.08	-.12
CC1 Online	-2.15	.79	-.23**	-2.18	.79	-.23**	-2.20	.79	-.23**	-2.21	.80	-.24**	-2.20	.80	-.24**	-1.93	.80	-.21*	-1.92	.80	-.21*
CC2 In-person	-.88	.38	-.21*	-.86	.39	-.21*	-.85	.39	-.20*	-.86	.39	-.21*	-.86	.39	-.21*	-.84	.39	-.20*	-.85	.39	-.20*
CC2 Online	.83	.90	.08	.74	.92	.07	.73	.92	.07	.75	.92	.07	.79	.93	.07	1.11	.93	.10	1.08	.94	.10
DIM				.19	.28	.06	.18	.28	.06	.19	.28	.06	.19	.28	.06	2.21	.88	.66*	2.19	.89	.66*
Threat							.13	.34	.03	.12	.35	.03	.11	.35	.03	5.42	2.21	1.29*	6.71	3.86	1.60
SIC Measure										-.04	.10	-.03	-.03	.10	-.03	.13	.39	.11	.49	.96	.41
SIC Manipulation													-.12	.35	-.03	-1.05	1.61	-.25	.30	3.67	.07
Thrt X SIC																-.21	.21	-.38	-.45	.63	-.82
Thrt X SIC Manip																.20	.69	.10	-.71	2.31	-.35
SIC X SIC Manip																.10	.20	.19	-.15	.63	-.28
DIM X Thrt																-1.33	.55	-1.28*	-1.32	.56	-1.26*
SIC X Thrt X SIC Manip																			.16	.40	.56
R^2		.20			.21			.21			.21			.21			.26			.26	
F		5.36			4.64			4.05			3.59			3.22			2.93			2.73	
p		<.001			<.001			<.001			<.001			.001			<.001			.001	

Notes. * $p < .05$
** $p < .01$

Regression analysis revealed no significant effect of stereotype threat when entered on step 3, (B = .13, SE = .34, p = .71). A review of the manipulation check showed that 62.9% of participants correctly identified the purpose of the study as explained in the threat manipulation they received, suggesting that the manipulation worked for these participants. Unexpectedly, a finding that emerged from this regression on step 6 was a significant interaction between DIM and the threat condition on test performance (B = -1.33, SE = .55, p = .02). The interaction revealed that those in the control/non-diagnostic condition demonstrated higher test scores as their identification with academics increased, although this did not contribute to the model ΔR^2= .05, ΔF (4,118) = 1.95, p = .11. In contrast, among participants in the threat/diagnostic condition, the observed relationship between DIM and test scores was negative (Figure 2). These results provide evidence that stereotype threat condition may have had a positive effect on performance among participants low in their identification with academics and less positive effect for those who identified more strongly.

Figure 2

Identification with Academics and Threat Manipulation Interaction on Test Scores.



Discussion

The results of study 2 do not support the hypothesis that social identity complexity can be changed and that having a more complex social identity will buffer the detrimental effects of stereotype threat. Drawing the self-concept maps had no effect on participants' social identity complexity, therefore providing no evidence that social identity complexity can be altered in the moment. Unlike the results of study 1, this study did not find any relationship between social identity complexity and test performance, providing no evidence that having a more complex social identity would buffer the detrimental effects of stereotype threat. The results of study 2 contradict the findings in Carmichael's (2011) dissertation research that found a positive relationship between SIC and test performance under stereotype threat.

As in study 1, there was a main effect of race, but this was confounded with the effect of survey location as the participants were not evenly distributed across locations. Due to the confounding variables, it is unclear if the racial differences indicate that the two racial groups may be too distinct to be included in a study together and should instead be studied separately, or if data collection across locations explains the variance.

These interaction results provide evidence that stereotype threat condition may have had a positive effect on performance among participants low in their identification with academics and less positive effect for those who identified more strongly. In contrast to previous stereotype threat research, this study revealed that participants in the threat/diagnostic condition outperformed those in the control/non-diagnostic condition. This finding suggests that the threat manipulation may have actually increased participants' motivation to do well on the task. Although this result was contrary to the stereotype threat effect expected, it is consistent with Jamieson and Harkins' (2007) research into the mere effort account that predicts that participants

faced with stereotype threat are motivated to do well, and if given the opportunity they well correct any known mistakes. The lack of the detrimental effects of stereotype threat as evidenced in Steele and Aronson's (1995) studies precludes this study from testing the hypothesis. It is possible that the stereotype threat manipulation was not effective as there was no main effect in the regression analysis and an examination of manipulation check results revealed that only 62.9% of participants correctly identified the purpose of the study as communicated to them. In either case the stereotype threat effect needed to test this hypothesis was not present so future research would need to successfully elicit the detrimental effects of stereotype threat to retest the intervention.

Another possible explanation for these results is that there is too much variation introduced in this study making the results unreliable, as in study 1. In both studies the participants were a mix of college and non-college enrolled adults, recruited through community colleges and an online survey platform. Approximately, half of the participants completed the survey outside of traditional academic settings, as many colleges moved to virtual means of delivery. This meant that participants were completing the study in countless different environments. All this variation may explain why there is no consistency in the findings between studies 1 and 2. In study 1 there was a relationship between SIC and test scores, although the data did not support the hypothesis, but in study 2 there was no statistically significant relationship found between SIC and test scores.

Perhaps that the reason these studies have not provided any evidence to support the hypothesis is that possessing a more complex social identity does not provide a buffering effect against stereotype threat. It is possible that among ethnic and racial minority group members possessing a more complex social identity does not allow one to protect against threats related to

their ethnic or racial identity as this identity is often salient and intertwined with many other social identities they possess. The social identity complexity measure used in this study is a measure of overlap complexity and cannot assess the salience of any specific identity. Further investigation and additional studies are needed to determine if there is sufficient support for this idea.

Limitations

There were three significant limitations of this study. As with study 1, the first limitation was the significant differences in test performance associated with survey location, making the racial differences uninterpretable. The additional variation from the inequitable distributions of both the different survey locations and two racial groups likely impacted the results of this study. In this study due to data collection, race and survey location are not independent variables and this could have significant implications on this analysis.

The second significant limitation of this study was in conducting it virtually and removing it from the relevant academic setting, as in study 1. It is possible that the stereotype threat effect did not have the expected effect because participants were removed from the stereotype relevant setting. In the meta-analysis conducted by Liu et al. (2021) all 181 studies included were conducted either in a lab or in the relevant field, unlike these dissertation studies that primarily took place online, removed from a lab or relevant field setting. Perhaps removing the stereotype relevant task from a more controlled setting such as a lab or the relevant environment reduces the threat and would contribute to lack of a main effect of stereotype threat.

The third limitation was that the manipulation check revealed that only 62.9% of participants correctly interpreted the threat condition to which they were assigned. This further

illustrates that the stereotype threat manipulation did not have the desired effect and without that threat, it is impossible to test an intervention designed to have a buffering effect.

CHAPTER 5

Study 3

Hypothesis

Study 3 tested the third and final hypothesis, that bolstered social identity complexity will increase the effectiveness of self-affirmation in improving performance under stereotype threat. Study 3 attempted to provide evidence for the previous hypotheses as the procedures in studies 1 and 2 are replicated, while adding one new component. Given the results of the two previous studies, confirming this hypothesis would provide the first evidence that supports the idea that there is a relationship between social identity complexity and academic performance when faced with stereotype threat and that possessing a more complex social identity buffers the detrimental effects of stereotype threat. Additionally, if this hypothesis were confirmed, it would provide evidence that social identity complexity can be changed and is not fixed. Lastly, confirming this hypothesis would not only confirm previous research findings, further confirming self-affirmation as an effective intervention to protect against stereotype threat, it would provide evidence that combining self-affirmation and bolstering social identity complexity through the creation of a self-concept map strengthens the buffering effect.

Methods

Participants

Based on power analysis using G*Power (Faul et al., 2007), a sample size of approximately 277 (approximately 35 in each group) will be necessary to obtain .8 power to detect a small effect (i.e., $f^2 = .04$). Given that this study is designed to measure the additive

effect of social identity complexity manipulation with the effect of self-affirmation, a small effect size was expected.

Participants included in study 3 were 310 adults (aged 18 and over) all recruited through Prolific. As in studies 1 and 2 the study design focused on college students, however, with the impact of COVID-19 making most college settings remote, the study recruitment was adjusted to include the online survey recruitment platform, Prolific. All study 3 participants were recruited through Prolific and were offered $6.75 and told the survey would take approximately 45 minutes of their time. As in study 2, participants recruited through Prolific were prescreened to allow the survey to only be viewed by respondents who were between the ages of 18-40, Nationality: United States, Ethnicity: Black/African American, Latino/Hispanic, having at least a high school diploma, and did not previously participate in study 1 or study 2.

Of the 310 respondents who completed the survey, 289 were included in the final analyses based on meeting the racial inclusion criteria of identifying as either African American/Black (71.6%) or Hispanic/Latinx (28.4%) and identifying as either male or female. The demographics of the 21 respondents removed from analyses were as follows: White/Caucasian (8), Other (2) Multi-racial (8), Black/African American (2), and Latinx/Hispanic (1). The three Black and Latinx participants removed from analysis identified as non-binary. Of the 289 participants included in the analyses, 48.8% identified as women and 51.2% identified as men. Most participants (65.1%) were not enrolled in a post-secondary institution when completing this survey. The age ranges of the participants recruited through Prolific are as follows: 18-24 (36.3%), 25-32 (55.4%), 33-40 (8.3%).

Design and Procedure

The study employed a 2 (threat condition: stereotype threat/diagnostic or control/non-diagnostic) X 2 (SIC manipulation: self-concept map or toy concept map) X 2 (self-affirmation

intervention: self-affirmation or control) experimental design. As in study 1 and 2, the survey was completed online for all participants using Qualtrics as the survey platform. Upon entering the survey, the first page displayed the informed consent form and participants had to check a box indicating their consent to continue to the rest of the survey.

After consenting, participants were then randomly assigned to a SIC manipulation, adapted from Gresky et al. (2005), as in study 2. Then participants were randomly assigned to receive the self-affirmation intervention or the control condition (Shapiro et al., 2012; Sherman & Cohen, 2006; Sherman et al., 2009).

The self-affirmation intervention is a task that requires participants to rank characteristics and values (Shapiro et al., 2012; Sherman & Cohen, 2006; Sherman et al., 2009). All participants received a list of 10 characteristics and values (e.g., sense of humor, artistic skills, and business/money) and ranked them in order of their importance to them, with a ranking of 1 being the most important and 10 as least important. Once participants completed the rankings, they were randomly assigned to either the self-affirmation intervention or control condition. Participants in the self-affirmation intervention condition wrote an explanation of why this value was important to them and provided one example of something they had done that demonstrated how important it was to them. Participants in the control condition wrote about their ninth ranked value with an explanation of why this value might be important to the typical American and provide one example of something the typical American might do to demonstrate how important it was to him or her.

Once completed, participants were then randomly assigned to the stereotype threat manipulation or control condition as in study 1 and 2. After receiving the threat condition, participants completed the verbal and math test, manipulation check, the social identity

complexity measure, and domain identification measure, also as in study 1 and 2. Then participants finished with responses to a demographics survey and were debriefed.

Measures

Academic test. Replicated from study 1 and study 2, the 10-item test consisted of five math and five verbal problems from the publicly available practice SAT. Reliability assessment revealed internal test-retest Cronbach alphas for both the math and verbal domains; $r(289) = .31$ and $r(289) = .61$, respectively. Collectively the 10-item test has a Cronbach's alpha of $\alpha = .59$. Although the coefficient is below the acceptable level of .7, this test was not intended to serve as a measure of intellectual ability, but instead a challenging task used to trigger and measure the stereotype effect.

Manipulation check. As in study 1 and 2, drawn from Steele and Aronson (1995), the manipulation check asked participants to select the purpose of the experiment:

The purpose of this experiment was to:

(e) Provide a genuine test of my mathematical and verbal abilities; or

(f) Provide a difficult assignment to examine factors involved in problem solving.

Social identity complexity measure. As in study 1 and 2, the social identity complexity measure was adapted from Miller et al. (2009), participants are probed for domains of social identities, including political organizations, gender identity, recreational and sports groups, religious affiliation, and ethnic/national identities, then asked to choose four that are most important to them (Appendix A).

Domain Identification Measure. Replicated from study 2, the Domain Identification Measure (Smith & White, 2001) was used to assess identification within the mathematics and English domains, as it was found to be related to the effectiveness of the social identity

intervention used in Gresky and colleagues' (2005) study. Reliability assessment revealed internal test-retest Cronbach alphas for both the math and English domains; $r(289)=.94$ and $r(289)=.87$, respectively. Collectively the 16-item test has a Cronbach's alpha of $\alpha = .84$.

Demographics and debriefing. As in study 2, the demographics section included race, gender, US residency status, postsecondary enrollment, grade level, highest educational level attainment, and age range. Once completed all participants were debriefed and thanked for their participation.

Results

Preliminary Analyses

Correlational analyses of the independent, dependent, and demographic variables revealed correlations between test scores and the SIC measure and race. SIC Score was negatively correlated with test scores ($r = -0.20$, $p < .001$). It is important to note that a low SIC score represents a more complex social identity, therefore the negative correlation with test scores indicates that participants with more complex social identities had higher test scores, providing tentative support for the hypothesis. Participants' race is positively correlated with test scores ($r = 0.28$, $p < .001$). T test analysis revealed that Latinx participants ($M = 6.74$, $SD = 1.97$) outperformed African American participants ($M = 5.36$, $SD = 2.18$) on the test, $t(287) = -5.01$, $p < .001$.

In addition to being correlated with test scores, race was correlated with SIC score and DIM. Race was negatively correlated with SIC score ($r = -0.15$, $p = .01$). The negative correlation with the SIC measure indicated that Latinx participants ($M = 5.06$, $SD = 1.30$) possessed more complex social identities than African American participants ($M = 5.55$, $SD = 1.49$), $t(287) = 2.57$, $p = .01$. Race was negatively correlated with DIM ($r = -0.15$, $p = .01$).

African American participants (M = 3.47, SD = .63) identified more strongly with academics than Latinx participants (M = 3.26, SD = .59), $t(287) = 2.60$, $p = .01$.

Main Analysis

To test the hypothesis that social identity complexity has an additive effect on test performance when done in conjunction with the existing stereotype threat intervention, self-affirmation, hierarchical regression analysis was employed. In the first step of the regression were the demographic variables: race and gender, followed by the DIM measure on step two. Step three was the threat manipulation, then the SIC measure on the fourth step, and on the fifth step the SIC manipulation. Then on the sixth step of the hierarchical regression the self-affirmation manipulation is introduced. The seventh step included six interactions variables: (1) threat manipulation and the SIC measure, (2) threat manipulation and SIC manipulation, (3) SIC measure and SIC manipulation, (4) DIM and threat manipulation, (5) threat manipulation and self-affirmation manipulation, (6) SIC manipulation and self-affirmation manipulation. The final step of the regression tested for the predicted 3-way interaction effect, where it is expected that participants who received the increased social identity exercise and the self-affirmation intervention will outperform those who were assigned to the respective control conditions when faced with stereotype threat. There is no predicted difference expected based on self-affirmation and SIC manipulation among those in the control/non-diagnostic condition.

Regression results indicate the final model explained 16.4% of variance (R^2 = .16, $F(14, 274) = 3.84$, $p < .001$). Results revealed no evidence of an additive effect of combining the SIC manipulation with the self-affirmation intervention as neither the two-way (B = -.48, SE = .50, p = .33) nor three-way interaction (B = .19, SE = 1.01, p = .85) proved statistically significant. These results are consistent with those of study 2 that found no effect of drawing the social

identity complexity maps. Unexpectedly, these results did not replicate previous findings in that there was no self-affirmation intervention effect (B = 0.31, SE = .25, p = .21) (Table 6).

Table 6

Hierarchical Multiple Regression Results Predicting Test Performance from Threat Condition, SIC Manipulation, and Self-affirmation

Predictor	Step 1			Step 2			Step 3			Step 4			Step 5		
	B	SE_B	β	*B*	SE_B	β	*B*	SE_B	β	*B*	SE_B	β	*B*	SE_B	β
Race	1.38	.28	.28**	1.50	.28	.31**	1.48	.28	.30	1.35	.28	.28**	1.37	.28	.28**
Gender	.16	.25	.04	.10	.25	.02	.08	.25	.02	.05	.25	.01	.05	.25	.01
DIM				.51	.20	.14*	.52	.20	.15	.58	.20	.17**	.58	.20	.16**
Threat							.22	.25	.05	.27	.25	.06	.27	.25	.06
SIC										-.28	.09	-.19**	-.28	.09	-.18**
SIC Manip													.24	.25	.05
Self-affirmation															
Thrt X SIC															
Thrt X SIC Manip															
SIC X SIC Manip															
DIM X Thrt															
Thrt X Self															
SIC Manip X Self															
SIC X Thrt X SIC Manip															
R^2		.08			.10			.10			.14			.14	
F		12.72			10.71			8.22			9.01			7.66	
p		<.001			<.001			<.001			<.001			<.001	

Notes. * $p < .05$
** $p < .01$

Table 6 Continued

Hierarchical Multiple Regression Results Predicting Test Performance from Threat Condition, SIC Manipulation, and Self-affirmation

Predictor	Step 6			Step 7			Step 8		
	B	SE_B	β	*B*	SE_B	β	*B*	SE_B	β
Race	1.36	.28	.28**	1.35	.28	.28**	1.35	.28	.28**
Gender	.04	.25	.01	.06	.25	.01	.06	.25	.01
DIM	.60	.20	.17**	.83	.65	.23	.83	.66	.24
Threat	.29	.25	.07	2.95	1.93	.67	3.42	3.13	.77
SIC	-.28	.09	-.18**	.08	.40	.05	.08	.40	.05
SIC Manip	.23	.25	.05	.22	1.40	.05	.63	2.56	.14
Self-affirmation	.31	.25	.07	1.57	1.13	.35	2.02	2.64	.46
Thrt X SIC				-.32	.18	-.54	-.32	.18	.18
Thrt X SIC Manip				.05	.50	.03	-.23	1.55	-.11
SIC X SIC Manip				.12	.18	.17	.12	.18	.18
DIM X Thrt				-.13	.40	-.11	-.14	.41	-.12
Thrt X Self				-.39	.50	-.19	-.68	1.64	-.38
SIC Manip X Self				-.48	.50	-.24	-.77	1.61	-.38
SIC X Thrt X SIC Manip							.19	1.01	.174
R^2		.15			.16			.16	
F		6.81			4.14			3.84	
p		<.001			<.001			<.001	

Notes. * $p < .05$
** $p < .01$

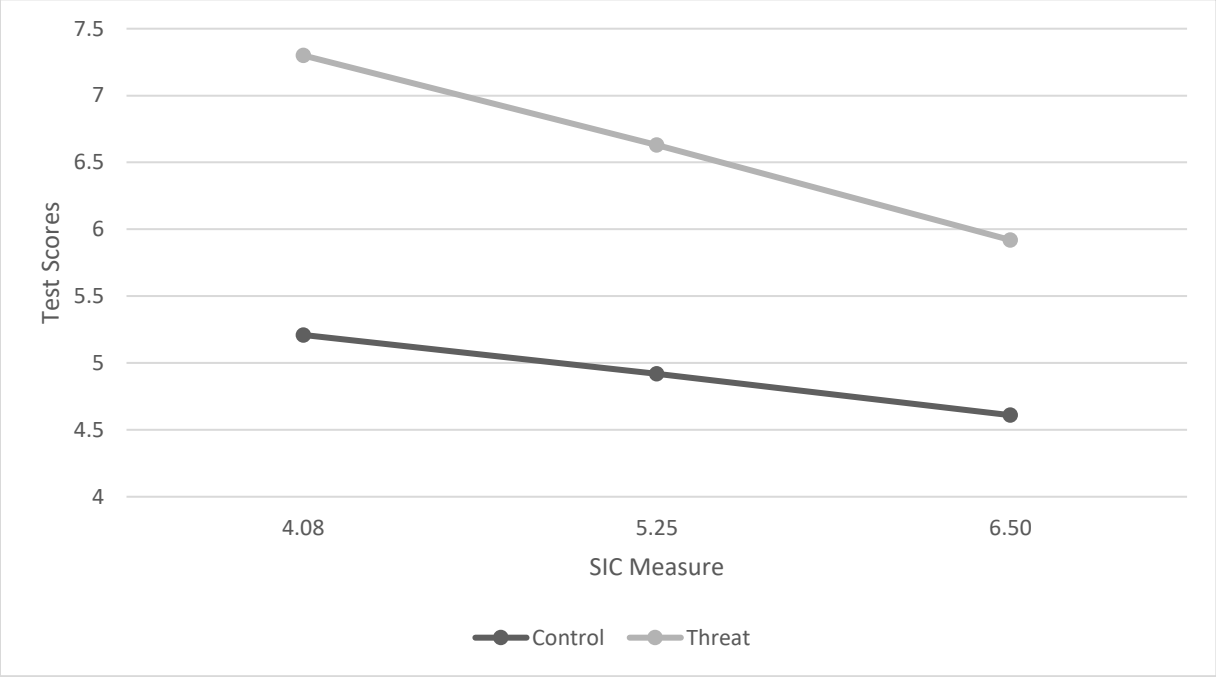
The regression results in step 4 revealed three significant main effects, SIC measure, race, and identification with academics, $\Delta R^2 = .03$, $\Delta F(1,283) = 10.99$, $p < .01$. SIC scores were predictive of test scores, in that those with more complex social identities earned higher test scores ($B = -0.28$, SE = .09, $p < .001$). This is consistent with the finding in study 1, where the interaction effect revealed that those in the control/non-diagnostic condition demonstrated higher test scores among those with higher social identity complexity. This finding provided additional evidence of the hypothesized relationship between SIC scores and test scores. In this study all participants were recruited via Prolific and so racial differences can be interpreted. Regression results revealed Latinx participants outperformed African American participants ($B = 1.38$, $SE = .28$, $p < .001$). Lastly, analysis revealed the DIM was a significant predictor of test scores ($B = 0.51$, SE = .20, $p = .01$) where those who more strongly identified with academics demonstrated higher test scores.

The only interaction to near statistical significance was the interaction which involved the threat manipulation and SIC measure ($B = -0.32$, $SE = .18$, $p = .070$). This interaction revealed that both participants in the threat/diagnostic condition and the control condition demonstrated higher test scores when they possessed a more complex social identity (Figure 3). Although both groups experienced an increase in test scores in relation to the possession of more complex social identities, this increase was greater among those in the threat condition. This interaction result is in complete contrast to the interaction in study 1. In study 1 SIC was related to test results only for those in the control/non-diagnostic condition, whereas in this study, SIC was significantly related with performance of both those in the threat/diagnostic and control/non-diagnostic condition. As in study 2, those in the stereotype threat condition outperformed those in the control condition, suggesting again that the threat manipulation may have increased participants'

motivation to do well on the test. This study showed that the stereotype threat effect was more positive for those high in SIC (low overlap) than for those low in SIC (high overlap).

Figure 3

SIC and Threat Manipulation Interaction on Test Scores.



To delve further into the SIC scores effect a ANCOVA was conducted to determine if the SIC manipulation was related to the SIC scores, while controlling for race and gender. The results showed that the SIC scores for the self-concept group (M = 5.29, SD = 1.23) and the toy concept group (M = 5.54, SD = 1.67) were not significantly different $F(1, 285) = 2.63$, $p = .11$. Therefore, the SIC manipulation did not significantly alter the complexity levels of participants' social identities, a finding that is consistent with the results of study 2. An unexpected finding that emerged from this analysis is that there was a difference in levels of social identity complexity based on race, $F(1, 288) = 6.59$, $p = .01$, such that Latinx participants (M = 5.06, SD = 1.30) demonstrated more complex social identities than African American participants (M = 5.54, SD = 1.49). This difference in SIC based on race was not found in study 2.

Discussion

The results of study 3 did not provide evidence to support the hypothesis that possessing a more complex social identity buffers against the detrimental effects of stereotype threat. The marginally significant interaction between SIC and stereotype threat suggest that the threat manipulation had a positive effect on participants and this effect was greater among those high in SIC. Due to the inconsistency of results between the three studies it is unclear if this result is evidence of the true relationship or simply Type 1 error. Type 1 error is postulated because the results in each of the three studies are inconsistent and contradictory. This lack of replicability and increased variability in survey administration indicates the effects found may be due to random chance. It is possible that the results of this study are more accurate than the previous two studies that had issues of confounded race and survey location. A replication of these studies without confounding variables is necessary to get a clearer understanding of these results.

An unexpected trend found in the interaction was that participants in the threat/diagnostic condition performed better than those in the control/non-diagnostic condition based on their level of social identity complexity. Although there was no main effect of the threat condition, this trend of participants in the threat condition outperforming those in the control condition suggests they may have benefited from an increased motivation to do well on the test as evidenced in Jamieson and Harkins' studies (2007). If the threat manipulation worked as expected the participants in the threat/diagnostic condition would have underperformed those in the control/non-diagnostic condition, with scores improving as their social identity complexity increased. It is recommended that this study is repeated with a detrimental threat effect to see if the interaction is replicable.

This study has failed to provide evidence that one's social identity can be manipulated through the creation of the self-concept maps, which is consistent with the findings of study 2. In both the current study and previous study 2 the SIC manipulation was unrelated to participants' SIC scores. The results found are likely based on the level of complexity the individuals possessed naturally and not a result of any manipulation from the study. This finding provides some support for the hypothesis that SIC buffers against the effects of stereotype threat, but there is no support that one's social identity complexity can be increased prior to the experience of the threat.

There is no evidence of the hypothesized additive effect of combining the activities designed to increase social identity complexity and self-affirmation. The regression results revealed no main effect or interaction involving self-affirmation, results that do not support the hypothesis of an additive effect. It is unclear why this study was not able to replicate previous research that have found significant evidence of self-affirmation as an intervention to help students combat the effects of stereotype threat. The exploration into the self-affirmation results in this study are explored further as part of the Ancillary Analysis.

In this study race was not confounded with survey location and differences associated with race were interpretable. In this study results revealed Latinx participants outperformed the African American participants on the test. This finding does suggest that the two groups may be heterogeneous in their academic performance and should not be consolidated in a study examining variations in test performance.

Finding that DIM is positively related to test performance, such that those who identify with academics do better on the test than those who do not identify with academics, is consistent with the findings in study 2 and previous research (Steele, 1997; Osborne & Walker, 2006).

In study 3, there was evidence that people with more complex social identities outperform those with less complex social identities when faced with stereotype threat; however, this was not found in the previous two studies. The inconsistency in findings across the three studies regarding the relationship between SIC and test performance under stereotype threat is the most puzzling outcome of these studies. All three studies displayed different relationships or no relationship at all. One explanation for this is that the findings in study 3 are due to Type 1 error and are occurring due to chance. The fact that a relationship between SIC and stereotype threat was found twice, suggests that this idea is worth further investigation. It is recommended that this study be conducted again, with an effective detrimental stereotype threat effect, with only African American students. Surveying only African American participants would reduce some of the variation attributed to race. Unfortunately, without a detrimental stereotype threat effect in this study, correlation analysis of the relationship between threat and performance within the African American sample would not allow for further examination of the hypothesis. The results of this proposed study would hopefully reveal the true nature of the relationship between SIC and stereotype threat.

Limitations

There were three significant limitations of this study. The first significant limitation of this study was in conducting it virtually and removing it from the relevant academic setting, as in study 1 and study 2. It is possible that the stereotype threat effect did not have the expected effect because participants were removed from the stereotype relevant setting. Although 73% of participants correctly answered the manipulation check, the lack of a main effect suggests that the manipulation did not elicit the impact expected. This is significant because this study was designed to test an intervention to a threat, without which there is no true test of the intervention.

The second limitation was the inability to manipulate SIC via the creation of self-concept maps. This manipulation did not work in study 2 or the current study. The failure to impact social identity complexity in the moment through the creation of self-concept maps does not allow for a test of the hypothesis, that bolstering social identity complexity before the experience of stereotype threat can buffer the detrimental effects of stereotype threat.

The third limitation of study was the inability to replicate the self-affirmation effect. This manipulation was adapted from previous literature where it was effective in protecting academic performance under threat. It is unclear why this intervention did not work in this study. Without the self-affirmation effect, there can be no test of an additive effect of another intervention.

CHAPTER 6

Ancillary Analysis

The idea at the center of this dissertation was based on Croizet and colleagues' (2001) reasoning, that one way to reduce the effect of stereotype threat is to increase the accessibility of other, non-threatened aspects of one's self-concept. It was theorized that this could be done using Brewer and Roccas' (2002) social identity complexity theory as the foundation to help African American and Latinx people increase the complexity of their social identities, thus increasing the accessibility of other, non-threatened aspects that could then allow them to buffer the effects of stereotype threat. Given the lack of evidence in the three studies of the hypothesized stereotype threat effect and inconsistent evidence of a relationship between SIC and test performance an ancillary analysis was conducted to further explore these hypotheses with less variation. This regression analysis included the threat condition, SIC measure, and interaction variable of the two as predictors of test scores among all African American participants across studies 1-3 (Table 7). Since gender was not included as a control variable in the ancillary analyses participants who identified as non-binary were included. By focusing solely on the African American participants, the analysis removes a significant amount of variation included in the previous studies due simply to race.

Table 7

Regression Results Predicting Test Performance from Threat Condition and SIC Measure

Predictor	*B*	SE_B	B
Threat	.98	.84	.22
SIC Measure	-.07	.11	-.05
SIC X Thrt	-.17	.15	-.24

Notes. R^2 =.02 (p = .10)

This regression analysis confirmed that there was no significant effect of stereotype threat or SIC on test performance among African American participants. Participants' test scores did not vary as a result of experiencing the threat condition or control condition, further confirming the findings in all three previous studies. This analysis also provided additional evidence that SIC was unrelated to test performance, confirming the results of study 2. Lastly, the regression analysis provided no evidence of an interaction effect between SIC and stereotype threat as was found in studies 1 and 3. These results confirmed that the stereotype threat manipulation did not elicit the hypothesized effect on test performance. The implications for the relationship between SIC and test performance are less clear but the results across the original three studies and this regression suggest the relationship in the previous studies may be the result of type 2 error.

Although the three studies previously detailed found no consistent evidence to support the idea that that one way to reduce the effect of stereotype threat is to increase the accessibility of other, non-threatened aspects of one's self-concept, there were important and interesting findings that emerged from the data that can contribute to our understanding of the social identities of racial/ethnic minorities. Due to the issues in studies 1 and 2 regarding confounding variables, race and survey location, ANCOVA analyses were conducted to control for survey location and allow for the exploration of racial differences.

Differences of Social Identity Complexity by Race

This dissertation specifically focuses on African American and Latinx participants because they are often the target of the detrimental effects of stereotype threat in academic settings. However, through the data collection process a large sample of participants were White and for this analysis were included. The question of interest was does social identity complexity vary by race? In study 3, there was a statistically significant difference in social identity

complexity between the African American and Latinx participants, so with the inclusion of data from the dominant racial group the question can be explored.

It is important to note that the data from all three studies were combined in this analysis. In all three studies participants completed the same Social Identity Complexity measure and so combining them to answer this question allows for a more robust sample.

Controlling for survey location, ANCOVA analysis revealed that there was a significant difference in social identity complexity scores between the three groups, $F(2, 814) = 3.57$, $p = .03$. The Bonferroni test revealed a significant difference in SIC scores between White and African American participants (Table 8). White participants possessed the most complex social identities, followed by Latinx and then African American participants. There was no significant difference in SIC scores based on survey location, $F(1, 814) = .55$, $p = .46$.

Table 8

ANCOVA: Differences in Social Identity by Race

	N	*M*	*SD*
Black/African American	347	5.48[a]	1.60
Latinx/ Hispanic	187	5.27	1.46
White/ Caucasian	284	5.17[a]	1.33

Notes. $F(2, 814) = 3.57$, $p = .03$)

[a] Indicates Bonferroni test of significant difference between groups

Racial majority members (White/ Caucasian) possessed a more complex social identity, as demonstrated by their lower mean estimates of overlap among ingroup memberships. These findings of racial differences in SIC scores are consistent with the findings of Brewer, Gonsalkorale and van Dommelen (2012) who found that Anglo-Australians perceived less

overlap among their ingroups than Asian-Australians. The findings in this data however add a complicating factor, the fact that the SIC scores of the Latinx group fell in between that of White and Black participants. For each comparison, the Latinx mean was not statistically different. This finding points to the heterogenous nature of the Latinx ethnic group. Latinx people can have a vastly different lived experiences and historical context with the dominant culture based on their skin color, country of ethnic origin, and family generations in America. It is hypothesized that this group's scores fall in middle because they are representative of varying perspectives that reflect the social identities of individuals who view themselves more similar to the dominant group and those that identify more as minority group members. Future research should be done that focuses specifically on the social identities of Latinx people and explore the variations attributed to the diversity of lived experiences, skin color, and generations in America.

Differences in Likelihood of Identifying Racial/Ethnic Group as First or Second Most Important Ingroup

An important aspect of the SIC measure was participants' selection of the four most important ingroups from the list of options. In contrast to the methodology utilized in study of Miller et al. (2009), where participants' race was automatically included in the pairings, participants in all three studies of this dissertation had the opportunity to select the four groups that were most important to them. This created an interesting opportunity for further exploration into the social identities of participants by race, the likelihood of identifying race/ethnicity as the first or second most important ingroup. This was of particular interest because of the racial difference in SIC scores. Differences between groups in how frequently they identify race/ethnicity as the most or second most important ingroup could aid in our understanding of

differences in the ways in which people organize their social identities as impacted by the group membership.

It is important to note again that the data from all three studies were combined in this analysis. In all three studies participants completed the same Social Identity Complexity measure and so combining them to answer this question allows for a more robust sample.

While controlling for survey location, ANCOVA analysis found that all three groups differed in the likelihood of selecting their race/ethnicity as the first or second most important ingroup ($F(2, 814) = 63.75$, $p < .001$). The Bonferroni test revealed significant differences between each group in their likeliness of indicating their race/ethnicity as most or second most important identity (Table 8). African Americans were significantly more likely to identify their race/ethnic ingroup as their most or second most important identity than Latinx or White participants. There was no significant difference in likelihood of selecting race/ethnicity as first or second most important ingroup based on survey location, $F(1, 814) = 1.75$, $p = .19$.

Table 8

ANCOVA: Differences in Selection of Race/Ethnicity as Most or Second Most Important Ingroup by Race

	N	*M*	*SD*
Black/African American	347	.78[a]	.41
Latinx/ Hispanic	187	.62[a]	.49
White/ Caucasian	284	.23[a]	.42

Notes. $F(2, 815) = 127.16$, $p < .001$

[a] Indicates Bonferroni test of significant difference between groups

Among African American participants 78% selected their race/ethnicity as the most or second most important ingroup, more than Latinx participants (62%) and White participants

(23%). This finding suggests that for African Americans their racial/ethnic identity is far more salient and possibly a dominant identity. Roccas and Brewer (2002) discussed four alternative structures of multiple ingroup representations, specifically Dominance, a way to organize one's group identities where one group identification is primary and all other potential group identities are subordinated. The findings of this analysis suggest that groups may differ in the structure of their social identities by race, and that for African Americans it is hypothesized that they possess a Dominance structure wherein their race is the primary identification.

Identification with Academics

Gresky and colleagues' (2005) found that creating self-concept maps as an intervention for women faced with stereotype threat in math experienced significantly better test scores when they were highly identified with math. Therefore, the measure of identification with academics through the Domain Identification Measure was essential in studies two and three when the SIC manipulation was introduced. Given the differences in test scores by race this measure is of significance because it is often a predictor of academic performance.

It is important to note that the data from studies 2 and 3 were combined in this analysis. In both studies participants completed the same Domain Identification measure and so combining them to answer this question allows for a more robust sample.

ANCOVA analysis revealed only a marginally significant difference between groups in their identification with academics while controlling for survey location, ($F(2, 605) = 2.92$, $p = .06$). The Bonferroni test revealed the marginally significant differences between African American and Latinx participants (Table 9). African Americans identified more strongly with academics than Latinx participants, with White participants in the middle. There was no

significant difference in identification with academics based on survey location, $F(1, 605) = 1.64$, $p = .20$.

Table 9

ANCOVA: Differences in Identification with Academics by Race

	N	*M*	*SD*
Black/African American	293	3.43	.63
Latinx/ Hispanic	134	3.29	.61
White/ Caucasian	182	3.36	.58

Notes. $F(2, 605) = 2.92$, $p = .06$

These results show that there is no significant or meaningful difference between groups in their identification with academics. Although there is a marginally significant difference between African American and Latinx participants, this difference is not meaningful and does not explain the differences in test performance. In study 3 Latinx participants outperformed African American participants, but their level of identification with academics was similar, with African Americans slightly more strongly identified. Steele's (1992) Disidentification theory explains how groups that are negatively stereotyped in academics, as African American and Latinx people often are, will remove the domain from their self-concept and instead focus on their identity on other domains that make their group positively distinct. In contrast to this theory African Americans and Latinx do not differ in their strength of identification with academics in comparison to their White counterparts. These results suggest that despite the differences in academic performance and experience of stereotype threat between the groups they all have similar levels of identification with academics as indicated by the DIM.

Differences in Effectiveness of Self-Affirmation

In study 3 the goal was to explore a possible additive effect of using two interventions to help participants buffer the detrimental effects of stereotype threat, thus the inclusion of the self-affirmation intervention. There is substantial evidence to the effectiveness of this intervention to improve academic performance under stereotype threat conditions (e.g., Cohen et al., 2006; Martens et al., 2006; Shapiro et al., 2013); however, in study 3 there was no effect of self-affirmation. This led to the question of the effectiveness of this intervention between the African American and Latinx participants. This question is of particular interest given the differences in social identity complexity and likeliness of identifying race/ethnicity as the most or second most important ingroup.

Since the self-affirmation intervention was only included in study 3 and the number of White participants was less than 10, the same sample as used in study 3 is included in the following regression analysis. Analysis of a possible difference in effectiveness of the self-affirmation intervention was done by conducting a regression with test scores as the dependent variable and DIM, race, gender, threat manipulation, self-affirmation intervention, and an interaction variable of race and self-affirmation as predictors.

Regression analysis revealed a significant self-affirmation effect and interaction effect on test scores ($R^2 = .13$, $F(7, 281) = 6.21$, $p < .001$) (Table 10). The interaction effect of self-affirmation and race indicate that the intervention was more effective in improving test scores for African American participants.

Table 10

Regression Self-Affirmation Effect on Test Scores Coefficients

	B	*SE B*	β
Gender	.05	.25	.01
DIM	.51	.20	.14*
Race	3.59	.84	.73**
Threat Manipulation	.26	.25	.06
SIC Manipulation	.32	.25	.07
Self-affirmation Manipulation	2.12	.74	.48**
Race X Self-affirmation	-1.43	.55	-.62**

Notes. $R^2 = .13$, $F(7, 281) = 6.21$, $p < .001$

* indicates significance at the .05 level

** indicates significance at the .01 level

A 2 (Race) X 2 (Self-affirmation intervention) ANOVA further illuminates the group differences, to reveal that the self-affirmation intervention helped African American participants to improve their test scores as predicted, but for Latinx participants their scores declined (Table 11).

Table 11

ANOVA: Test scores by self-affirmation and race

		Self-Affirmation	Control
	N	*M*	*M*
Black/African American	207	5.71	5.07
Latinx/ Hispanic	82	6.31	7.14

Notes. $F(1, 285) = 7.18$, $p = .01$)

The result of this analysis reflects two very different effects of the self-affirmation intervention, one in which African Americans benefited and another in which Latinx participants' performance worsened. This finding provides greater clarity into the results of study

3, as it suggests that the self-affirmation intervention did work as expected for the African American participants. The results for the Latinx participants are consistent with those of Voisin et al. (2019), who found evidence that self-affirmation can decrease performance under stereotype threat. Their research revealed that self-affirmation decreases the motivation to disconfirm the negative stereotype because self-integrity is reinforced, which then reduces motivation in a stereotype threat situation. The risk of failure may not represent a threat and the internalized stereotype may be less relevant to self-integrity. Furthermore, this result shows that once again Latinx participants have a different experience as it relates to their identity. It is unclear why the self-affirmation intervention had contrasting effects on test performance for African American and Latinx participants. It is possible that Latinx participants were affirmed and this removed the threat and reduced their motivation, whereas for African American participants the self-affirmation intervention helped them to buffer the detrimental effects without reducing their motivation. Future research should be done to better understand racial differences in the experience of self-affirmation under stereotype threat.

CHAPTER 7

General Discussion

The results of all three studies provided inconsistent evidence and therefore did not reliably support the general hypothesis based on Croizet and colleagues' (2001) reasoning that one way to reduce stereotype threat is to increase the accessibility of other, non-threatened aspects of one's self-concept by increasing one's social identity complexity. It is hypothesized that the inconsistency in results across the three studies is partly due to the inclusion of too much variability. Having to move the study procedures outside of the classroom due to COVID-19 meant that participants were in not in the traditional educational setting and this may have reduced the effect of stereotype threat.

A meta-analysis of the effectiveness of stereotype threat interventions conducted by Liu et al. (2021) found that stereotype threat interventions conducted in a lab were more effective than those conducted in the field (i.e., grade/professional schools and in continuous education/training programs). This meta-analysis reviewed 181 stereotype threat intervention studies and all of them were conducted either in a lab or in the relevant field, unlike these dissertation studies that primarily took place online, removed from a lab or relevant field setting. According to Steele et al. (2002) stereotype threat is a situational threat that arises from situational cues signaling that a negative stereotype about one's social identities is now relevant as a possible interpretation for one's behavior and self in the setting. Although the task description should have been sufficient to activate the detrimental stereotype threat effect, it is possible that this virtual setting, removed from the educational context was not sufficient to elicit the threat of being judged according to the negative stereotype. Therefore, taken together the fact that stereotype threat is a situational threat and that interventions conducted in a lab setting are

more effective than those in the field, and these studies were further removed from the relevant context it stands to reason that the intervention and the stereotype threat effect itself were ineffective due to the setting.

Study 1 sought initial evidence for the buffering effect of social identity complexity on test performance under stereotype threat. The results revealed a relationship between SIC scores and test performance where those in the control condition demonstrated higher test scores when they possessed more complex social identities, but among those faced with stereotype threat there was no relationship between SIC and test performance. This finding was the first sign that the hypothesized relationship may not exist as expected, however, there is evidence that people do vary on a continuum from simple to complex in their organizations of their social identities as theorized by Roccas and Brewer (2002). Although the relationship revealed did not support the hypothesis, it did show evidence that there is a positive relationship between social identity complexity and academic performance among those in the control condition.

Study 2 was designed build on the previous study, and experimentally test for a possible buffering effect against stereotype threat using an intervention designed to increase social identity complexity. Again, the results did not support the hypothesis, as there was no effect of the SIC manipulation on test scores or SIC scores. Roccas and Brewer (2002) theorized that social identity complexity varies across individuals and across situations, though unfortunately, drawing self-concept maps did not elicit any situational change in social identity complexity. These results also did not provide evidence to support the findings of Gresky et al. (2005) who used self-concept maps to buffer the effects of stereotype threat among women in math.

Lastly, study 3 explored a possible additive effect of social identity complexity on test performance, when combined with self-affirmation. This study demonstrated the first evidence

of the hypothesized relationship between social identity complexity and test scores. Participants in the stereotype threat and control conditions demonstrated higher test scores when they possessed more complex social identities. However, the results revealed no evidence of an effect of the SIC manipulation or self-affirmation intervention. Ancillary analyses revealed that self-affirmation intervention was effective but only for the African American sample, thus providing support for previous research on the effectiveness of the intervention (e.g., Cohen et al., 2006; Martens et al., 2006; Shapiro et al., 2013).

Although these studies failed to produce consistent evidence of the hypothesized relationship between SIC and academic performance under stereotype threat, ancillary analyses provide some paradigm altering findings that call the general idea at the center of these studies into question. The general idea at the center of this dissertation is based on the theory that one can reduce stereotype threat by increasing the accessibility of other, non-threatened aspects of one's self-concept. This idea is likely impossible, particularly for African Americans who perceive more overlap in their ingroups, therefore possessing a less complex social identity than their White counterparts. Additionally, African Americans were far more likely to identify their race/ethnicity as the most or second most important ingroup, thus possibly holding racial/ethnic identity as their primary identification, to which other identities are subordinate. These findings are consistent with Roccas and Brewer's (2002) theory that those facing an ingroup threat are likely to increase the salience of the threatened ingroup, thus temporarily dominating social identity and membership in other ingroups become both less important and less differentiated. Additionally, according to Brewer et al. (2012) ethnic minorities are more likely to "carry" their ethnic identity across social contexts, so that their membership in other social groups and categories is subjectively linked or associated with their ethnic identity. This means that for most

African Americans it would be impossible to disassociate from their racial/ethnic identity when under threat because it is likely inextricably connected to their other identities. African American people are members of many other ingroups, but often their racial/ethnic identity is still linked, such that they are not just "scholars" but "Black scholars," not just "women" but "Black women."

The results for the Latinx participants are less clear as 62% of Latinx participants identified their race/ethnicity ingroup as most or second most important, in comparison to 78% among African Americans. Additionally, Latinx SIC scores fell in between those of White and African American participants. It is hypothesized that Latinx people would fall into two groups, one that is more similar to African Americans in their conceptions of their identity, and another that would be more similar to White people, and the determining factors would be based on skin color, country of ethnic origin, and number of family generations in America.

Lastly, this study has found substantial evidence that African American and Latinx populations are not similar in their conceptions of identity and therefore should be tested separately for any test of interventions designed to impact their self-concept. Originally these two groups were included together for this dissertation because they had similar experiences of stereotype threat. However, the results of these studies have shown that they significantly differ in test scores, likeliness of holding their race/ethnicity as their most or second most important ingroup, and the effectiveness of the self-affirmation intervention. Although these groups share many similarities in their lived experiences, the variances between them are meaningful and each deserves to be studied independently. Again, it is recommended that any studies exploring the identities of Latinx people consider background characteristics that are likely significantly related to their self-concepts.

Limitations

The main limitation of these studies was the inability to replicate the main effect of stereotype threat that is possibly attributed to the setting in which the study was conducted. Stereotype threat effects have been replicated in countless studies, including this researcher's master's thesis study that was conducted in high schools. Conducting these studies so far removed from the relevant context of the educational setting has likely contributed to the inconsistencies across studies.

Another critical limitation was the unintended consequence of data collection for studies 1 and 2 that resulted in the confounding race and survey location variables. It is likely that the results of both studies are comprise and unreliable as a result.

Another limitation of these studies was the inclusion of African American and Latinx participants, without the intention to compare the two groups. These two groups are distinct in their racial/ethnic identities and these differences impact their experience of stereotype threat in distinct ways. Having both groups introduced more unnecessary variation to these studies.

The inability to replicate the self-affirmation effect eliminated any opportunity to test for an additive effect. It is still unclear why there was a racial difference in the experience of self-affirmation, but this resulted in the inability to test the third hypothesis.

Future Research

Future research should focus on understanding more about identity structures, as proposed by Roccas and Brewer (2002) and the implications of each. Roccas and Brewer (2002) introduced four structures, and research should be conducted to determine if racial minorities are more likely to hold any of four structures and what that means for their lived experiences. If African Americans are more likely to possess a dominance social identity structure it is unlikely

any intervention designed to increase SIC would work to alleviate the detrimental effects of stereotype threat. Furthermore, understanding the different SIC structures people of different racial identities possess would help to tailor new identity threat interventions to their specific needs and increase the chances that the intervention would have a positive impact. There is now more evidence that shows racial/ethnic minority group members have less complex social identities than White people, but research must go further to better understand how these social identity structures work for different groups.

Future research should explore different testing situations to ascertain a threshold for when stereotype threat effect is beneficial and then when it becomes detrimental. This research should be conducted with a relevant academic task, one stereotyped racial group, in a lab or applicable educational setting. This study should include two manipulations of threat (explicit and implicit) and two testing environments (one that is timed and one that is not) to examine which combination yields beneficial and detrimental stereotype threat effects.

Another proposed study could test a new intervention designed to increase effort without increasing cognitive load to result in the mere effort effect. Black and White participants could be randomly assigned to one of three conditions before completing a challenging academic task: 1) told that the test has been found to produce racial differences, 2) told that the test has been found to produce racial differences, but research has shown that those who check their work before submission outperform those who do not, or 3) told that the test is a measure of problem-solving skills. This would allow for a comparison of an intervention designed to elevate threatened participants' motivation to perform well without increasing cognitive load and decreasing performance, thus resulting in the beneficial mere effort effect. By including both a stereotype threat targeted group and a non-threatened group the study could have multiple

opportunities to compare between and within groups. This potential study could provide another simple intervention that could be used to help stereotype targeted groups turn a detrimental effect into a beneficial one.

Conclusion

This dissertation was designed to test the idea that African American and Latinx people could protect against stereotype threat by increasing the availability of salient non-threatened identities, resulting in a potential intervention that could be used in classrooms everywhere to help students overcome the detrimental effects of stereotype threat. Although these studies failed to provide consistent support for this idea and intervention, important findings emerged regarding the distinctions in social identities between African American and Latinx that are essential to progress the study of identity. Too often studies regarding identity focus on the dominant group and from this group theories are derived and assumed to apply to everyone. This research revealed important distinctions in the social identities of participants that were based on race. This dissertation may not have found evidence for an intervention to help African American and Latinx students buffer the effects of stereotype threat, but perhaps because of the results showing the importance of considering race/ethnicity in identity research more interventions will be found that specifically address the unique identity threat these two groups experience.

Reference List

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*(1), 29-46.

Beilock, S. L. (2008). Math performance in stressful situations. *Current Directions in Psychological Science, 17*(5), 339-343.

Ben-Zeev, T., Fein, S., & Inzlicht, M. (2005). Arousal and stereotype threat. *Journal of Experimental Social Psychology, 41*(2), 174-181.

Brewer, M. B. (2008). Deprovincialization: Social identity complexity and outgroup acceptance. In U. Wagner, L. R. Tropp, G. Finchilescu, & C. Tredoux (Eds.) *Improving intergroup relations: Building on the legacy of Thomas F. Pettigrew*, (pp. 160-176). Malden: Blackwell.

Brewer, M. B., Gonsalkorale, K., & van Dommelen, A. (2012). Social identity complexity: Comparing majority and minority ethnic group members in a multicultural society. *Group Processes & Intergroup Relations, 16*(5), 529-544.

Brewer, M. B., & Pierce, K. P. (2005). Social identity complexity and outgroup tolerance. *Personality and Social Psychology Bulletin, 31*(3), 428-437.

Carmichael, C. (2011). *Stereotype threat: Threat accrual and social identity complexity* (Doctoral dissertation). Retrieved from ProQuest. (Accession Order No. AAI3459574)

College Board. (2015). *2015 College-Bound Seniors: Total Group Profile Report.*

Croizet, J.-C., Désert, M., Dutrévis, M., & Leyens, J.-P. (2001). Stereotype threat, social class, gender, and academic under-achievement: When our reputation catches up to us and takes over. *Social Psychology of Education, 4*(3-4), 295-310.

Désert, M., Leyens, J.-P., Croizet, J.-C., & Klopfenstein, T. (2001). *Reducing stereotype threat through re-individuation.* Manuscript submitted for publication.

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

Faul, F., Erdfelder, E., Lang A. G., & Buchner A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Forbes, C. E., Schmader, T., & Allen, J. J. B. (2008). The role of devaluing and discounting in performance monitoring: A neurophysiological study of minorities under threat. *Social Cognitive Affective Neuroscience, 3*(3), 253-261.

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*(5), 659-670.

Gresky, D. M., Ten Eyck, L. L., Lord, C. G., & McIntyre, R. B. (2005). Effects of salient multiple identities on women's performance under mathematics stereotype threat. *Sex Roles, 53*, 703-716.

Harris, L. (2010). Black racial identity and academic Success: The effect of stereotype threat on urban high school students (Unpublished master's thesis). Claremont Graduate University, Claremont, CA.

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

Inzlicht, M., & Schmader, T. (2012). *Stereotype threat: Theory, process, and application.* New York, NY: Oxford University Press, Inc.

Jamieson, J. P., & Harkins, S. G. (2007). Mere effort and stereotype threat performance effects. *Journal of Personality and Social Psychology, 93*(4), 544-564.

Liu, S., Liu, P., Wang, M., & Zhang, B. (2021). Effectiveness of stereotype threat interventions: A meta-analytic review. *Journal of Applied Psychology, 106*(6), 921-949.

Markus, H. & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology, 38,* 299–337.

Miller, K.P., Brewer, M.B., & Arbuckle, N.L. (2009). Social identity complexity: Its correlates and antecedents. *Group Processes and Intergroup Relations, 12,* 79-94.

Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control.* New York: Cambridge University Press.

Nadler, J. T. & Clark, M. H. (2011). Stereotype Threat: A Meta-Analysis Comparing African Americans to Hispanic Americans. *Journal of Applied Social Psychology, 41*(4), 872-890.

Osborne, J. W. & Walker, C. (2006) Stereotype threat, identification with academics, and withdrawal from school: Why the most successful students of colour might be most likely to withdraw. *Educational Psychology*, *26*(4), 563-577.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology, 70,* 153-163.

Roccas, S., & Brewer, M. B. (2002). Social identity complexity. *Personality and Social Psychology Review, 6*(2), 88-106.

Schmader, T., & Beilock, S. (2012). An integration of processes that underlie stereotype threat. In M. Inzlicht & T. Schmader (Eds.), *Stereotype threat: Theory, process, and application*

(33-50). New York, NY: Oxford University Press, Inc.

Schmader, T., Forbes, C. E., Zhang, S., & Mendes, W. B. (2009). A meta-cognitive perspective on cognitive deficits experienced in intellectually threatening environments. *Personality and Social Psychology Bulletin, 35*(5), 584-596.

Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85*(3), 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.

Schmid, K., & M. Hewstone. (2011). Social identity complexity: Theoretical implications for the social psychology of intergroup relations. In R. M. Kramer, G. J. Leonardelli, & R. W. Livingston (Eds.) *Social Cognition, Social Identity, and Intergroup Relations: A Festschrift in Honor of Marilynn B. Brewer (pp. 77-102).* Psychology Press: New York, NY.

Shapiro, J. R., & Neuberg, S. L. (2007). From stereotype threat to stereotype threats: Implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review, 11*(2), 107–130.

Shapiro, J. R., Williams, A. M., & Hambarchyan, M. (2013). Are all interventions created equal? A multi-threat approach to tailoring stereotype threat interventions. *Journal of Personality and Social Psychology, 104*(2), 277-288.

Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: Self-affirmation theory. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology, 38,* 183- 242. San Diego, CA: Academic Press.

Sherman, D. K., Cohen, G. L., Nelson, L. D., Nussbaum, A. D., Bunyan, D. P., & Garcia, J. P. (2009). Affirmed yet unaware: Exploring the role of awareness in the process of self-affirmation. *Journal of Personality and Social Psychology, 97*(5), 745–764.

Smith, J. L., & White, P. H. (2001). Development of the domain identification measure: A tool for investigating stereotype threat effects. *Educational and Psychological Measurement, 61*(6), 1040-1057.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4-28.

Steele, C. M. (1992). Race and the schooling of black Americans. *The Atlantic Monthly, 4,* 68-78.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), *613-629.*

Steele, C. M. (2010). *Whistling Vivaldi: And Other Clues to How Stereotypes Affect Us (Issues of Our Time).* W. W. Norton & Company.

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.

Steele C. M., Spencer S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology, Vol. 34* (379-440). New York, NY: Academic Press.

U.S. Department of Education, National Center for Education Statistics. (2019). Digest of Education Statistics, 2018 (NCES 2020-009), Table 226.10. https://nces.ed.gov/fastfacts/display.asp?id=171

Voisin, D., Brick, C., Vallée, B., & Pascual, A. (2019). When stereotype threat does not impair

performance, self-affirmation can be harmful. *Self and Identity, 18*(3), 331-348.

Whitley, B. E., Jr., & Kite, M. E. (2006). The psychology of prejudice and discrimination. Belmont, CA: Thomas Higher Education.

Appendix A
Social Identity Complexity Measure

Think about the different groups you belong to and come up with the four that are most important to you. Below are the possible responses.

Is your race/ethnicity group important to you?

Black/African American, White/Caucasian, Hispanic/Latino, Asian American/Pacific Islander, Native American, Multi-racial

Is your gender important to you?

Female, Male, Transgender

Is your religious affiliation important to you?

Protestant (Christian, non-Catholic), Catholic, Jewish, Muslim, Atheist (no religious affiliation)

Is your nationality or local pride important to you?

American, Local Identity (i.e., Ohioan, Southerner, New Yorker, Texan, Chicagoan)

Is your occupational identity important to you?

Student, Entrepreneur, Occupational Identity (i.e., Educator, Healthcare professional, Business professional, IT Tech)

Is your political affiliation important to you?

Democrat, Republican, Independent

Is your political worldview important to you?

Liberal, Conservative, Moderate

Is a particular interest group important to you?

Sorority/fraternity member, scholar, athlete, scientist, musician, writer, artist

1. What is the most important group you belong to? (Select only one.)

Black/African American	Jewish	Conservative
White/Caucasian	Muslim	Moderate
Hispanic/Latino	Atheist (do not believe in a god)	Scholar
Asian American/Pacific Islander	American	Activist
Native American	Local Identity (ie. Ohioan, Southerner, New Yorker, Texan, Chicagoan)	Athlete
Multi-Racial	Sorority/Fraternity Member	Scientist
Female	Student	Writer
Male	Democrat	Musician

Non-binary	Republican	Entrepreneur
Protestant (Christian, non-Catholic)	Independent	Occupational Identity (ie. Educator, Healthcare professional, Business professional, IT Tech)
Catholic	Liberal	Artist

2. What is the second most important group you belong to? (Select only one.)

Black/African American	Jewish	Conservative
White/Caucasian	Muslim	Moderate
Hispanic/Latino	Atheist (do not believe in a god)	Scholar
Asian American/Pacific Islander	American	Activist
Native American	Local Identity (ie. Ohioan, Southerner, New Yorker, Texan, Chicagoan)	Athlete
Multi-Racial	Sorority/Fraternity Member	Scientist
Female	Student	Writer
Male	Democrat	Musician
Non-binary	Republican	Entrepreneur
Protestant (Christian, non-Catholic)	Independent	Occupational Identity (ie. Educator, Healthcare professional, Business professional, IT Tech)
Catholic	Liberal	Artist

3. What is the third most important group you belong to? (Select only one.)

Black/African American	Jewish	Conservative
White/Caucasian	Muslim	Moderate
Hispanic/Latino	Atheist (do not believe in a god)	Scholar
Asian American/Pacific Islander	American	Activist

Native American	Local Identity (ie. Ohioan, Southerner, New Yorker, Texan, Chicagoan)	Athlete
Multi-Racial	Sorority/Fraternity Member	Scientist
Female	Student	Writer
Male	Democrat	Musician
Non-binary	Republican	Entrepreneur
Protestant (Christian, non-Catholic)	Independent	Occupational Identity (ie. Educator, Healthcare professional, Business professional, IT Tech)
Catholic	Liberal	Artist

4. What is the fourth most important group you belong to? (Select only one.)

Black/African American	Jewish	Conservative
White/Caucasian	Muslim	Moderate
Hispanic/Latino	Atheist (do not believe in a god)	Scholar
Asian American/Pacific Islander	American	Activist
Native American	Local Identity (ie. Ohioan, Southerner, New Yorker, Texan, Chicagoan)	Athlete
Multi-Racial	Sorority/Fraternity Member	Scientist
Female	Student	Writer
Male	Democrat	Musician
Non-binary	Republican	Entrepreneur
Protestant (Christian, non-Catholic)	Independent	Occupational Identity (ie. Educator, Healthcare professional, Business professional, IT Tech)
Catholic	Liberal	Artist

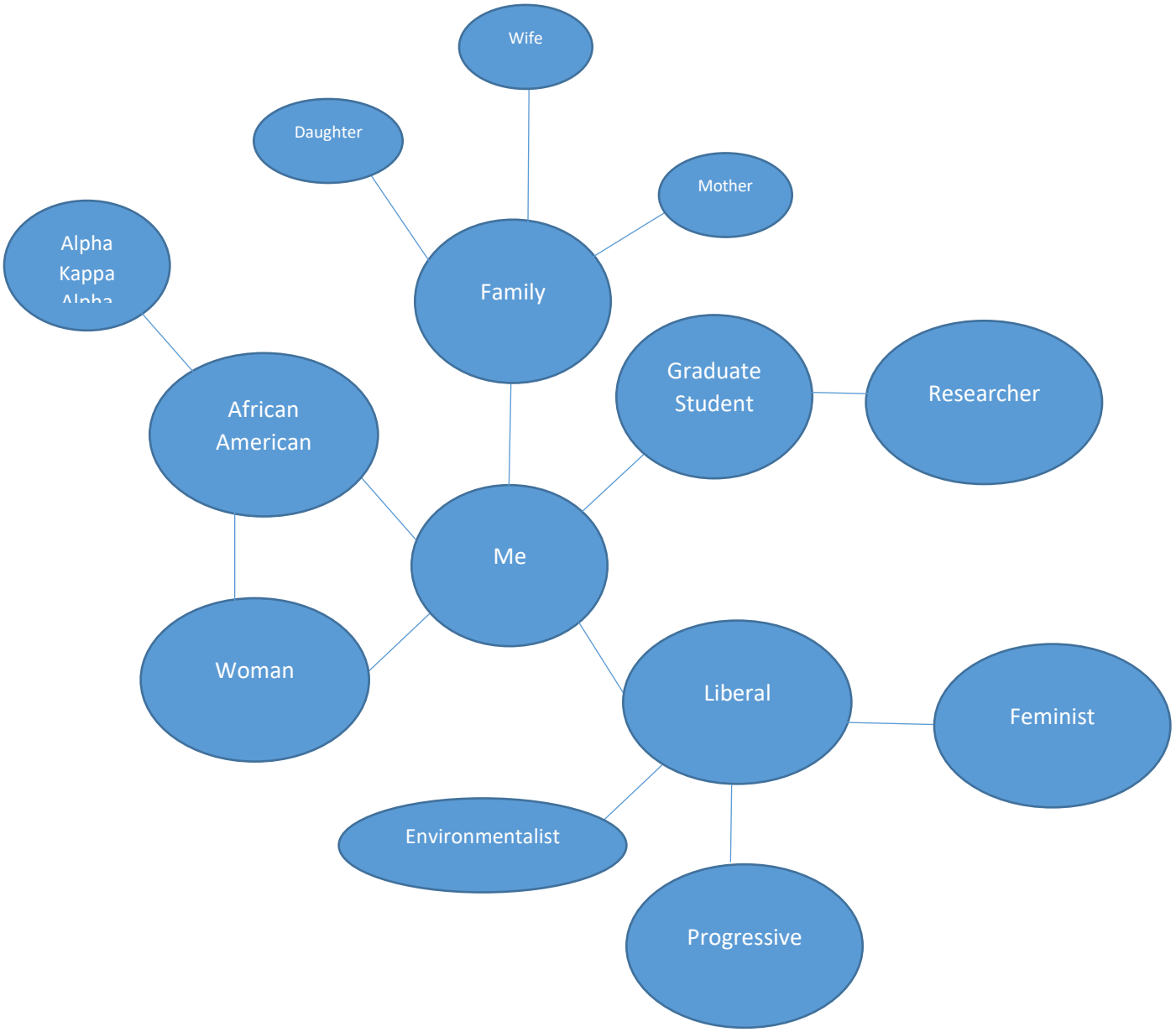
5. Sometimes members of one group also belong to other groups. I'd like you to rate how much the membership of the different groups overlaps on a scale from 0 to 10. If no members of the first group are also members of the second group, then rate the overlap as 0. If about half of the first group are also members of the second group, then rate the overlap as 5. And if all of the members of the first group are also members of the second group, then rate the overlap as 10. You can use any number from 0 to 10 to rate the amount of overlap between the two groups as you think about them.

	None are			Half are					All are		
	0	1	2	3	4	5	6	7	8	9	10
How many (answer to #1) are also (answer to #2)?											
How many (answer to #1) are also (answer to #3)?											
How many (answer to #1) are also (answer to #4)?											
How many (answer to #2) are also (answer to #1)?											
How many (answer to #2) are also (answer to #3)?											
How many (answer to #2) are also (answer to #4)?											
How many (answer to #3) are also (answer to #1)?											
How many (answer to #3) are also (answer to #2)?											
How many (answer to #4) are also (answer to #4)?											
How many (answer to #4) are also (answer to #1)?											
How many (answer to #4) are also (answer to #2)?											
How many (answer to #4) are also (answer to #3)?											

Appendix B

Increased Social Identity Complexity Condition

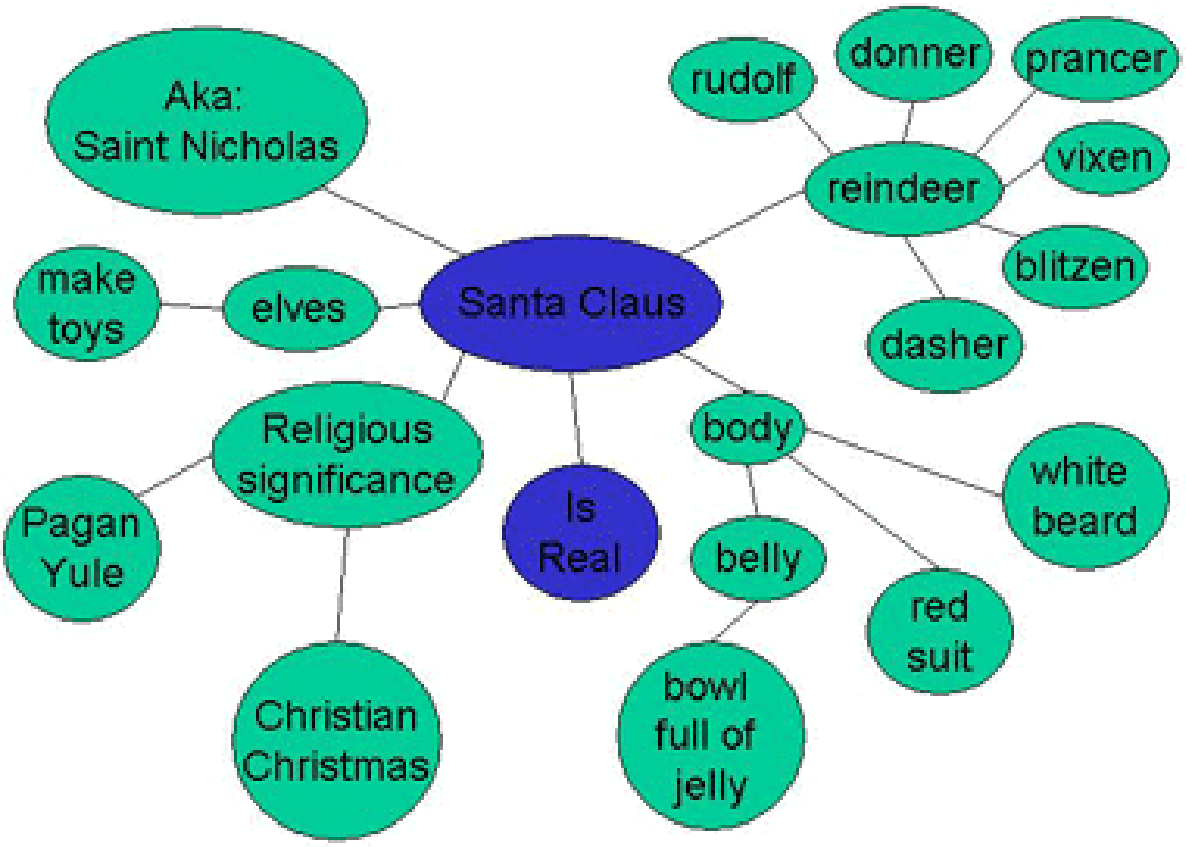
Please draw a self-concept map that represents your different social identities. At the core of the self-concept map will be "me" and then you will draw lines connecting to other identities. An example is provided below. Try to create a self-concept map with plenty of information, using multiple connections and many nodes (e.g., school, family, interests).



Appendix C

Increased Social Identity Complexity Control Condition

Please draw a self-concept map that represents your thoughts and knowledge about toys. At the core of the concept map will be "toys" and then you will draw lines connecting to other related ideas. An example is provided below. Try to create a concept map with plenty of information, using multiple connections and many nodes (e.g., board games, dolls, blocks, balls).



Appendix D

Range, Mean Scores, Standard Deviations, and Correlations for Sociodemographic Variables, Independent, and Dependent Variables

		Range	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	Test Score	0-9	4.58	2.10																
2	Race	1, 2	1.39	.49	.30**															
3	Gender	1, 2	1.42	.50	.09	-.05														
4	CC1 In-person	0, 1	.04	.19	-.07	.09	-.09													
5	CC1 Online	0, 1	.05	.22	-.19*	-.05	-.06	-.04												
6	CC2 In-person	0, 1	.49	.50	-.21*	-.37**	.06	-.19*	-.23**											
7	CC2 Online	0, 1	.04	.19	.13	.01	-.01	-.04	-.04	-.19*										
8	Prolific	0, 1	.39	.49	.28**	.37**	.01	-.15	-.18	-.78	-.15									
9	DIM	1.92-10	3.35	.63	.08	-.03	.20*	.04	.05	-.11	.16	.02								
10	Stereotype Threat Mnp	0, 1	1.52	.50	-.05	-.18*	.01	.03	.09	-.03	.03	-.03	.07							
11	SIC Measure	1.56- 5	5.48	1.74	.01	.05	-.07	.11	-.02	-.10	.07	.04	.12	-.04						
12	SIC Manipulation	0, 1	1.48	.50	-.02	-.03	.08	-.03	-.02	-.00	.12	-.03	.04	-.06	.09					
13	Thrt X SIC Score	0-10	8.29	3.77	-.05	-.11	-.01	.05	.04	-.07	.08	.01	.13	.70**	.65**	-.01				
14	Thrt X SIC Manip	0-1	2.24	1.06	-.05	-.16	.07	-.01	.04	-.03	.10	-.02	.07	.66**	.01	.68**	.47**			
15	SIC Score X SIC Manip	0-9.33	8.17	4.02	.01	.03	.02	.07	-.05	-.08	.15	.02	.09	-.09	.70**	.74**	.40**	.45**		
16	Id w/ Academics X Thrt	0-5	4.86	2.27	-.00	-.17*	.06	-.05	.13	-.07	.13	-.02	.54**	.72**	.03	-.07	.55**	.44**	-.06	
17	SIC Measure X Threat X SIC Manip	0-1	12.25	7.13	-.04	-.11	.05	.05	.00	-.08	.15	.01	.10	.50**	.55**	.58**	.77**	.79**	.77**	.36**

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

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

Appendix E

Range, Mean Scores, Standard Deviations, and Correlations for Sociodemographic Variables, Independent, and Dependent Variables

		Range	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Test Score	0-10	5.75	2.21														
2	Race	1, 2	1.28	0.45	.28**													
3	Gender	1, 2	1.51	.50	.04	.02												
4	DIM	1.56- 5	3.41	.62	.10	-.15**	.10											
5	Stereotype Threat Mnp	1, 2	1.53	.50	.06	.07	.08	-.09										
6	SIC Measure	1.92- 10	5.41	1.45	-.21**	-.15*	-.02	.11	.04									
7	SIC Manipulation	1, 2	1.54	.50	.06	-.06	.01	.05	-.03	-.09								
8	Self-Affirmation	1, 2	1.46	.50	.06	.02	.02	-.08	-.04	.02	.02							
9	Thrt X SIC Score	2-20	8.30	3.71	-.11	-.05	.05	.01	.75**	.66**	-.08	-.04						
10	Thrt X SIC Manip	1-4	2.35	1.10	.08	.01	.06	-.05	.68**	-.04	.68**	-.02	.47**					
11	SIC Score X SIC Manip	2.67-18.67	8.28	3.30	-.07	-.13	-.01	.09	.00	.54**	.76**	.06	.34**	.53**				
12	DIM X Thrt	1.56- 10	5.19	1.91	.10	.00	.14	.43**	.85**	.10	-.02	-.08	.69**	.58**	.04			
13	Thrt X Self-affirm	1- 4	.23	.42	.06	.07	.08	-.11	.52**	-.04	-.02	.60**	.35**	.35**	-.01	.40**		
14	SIC manip X Self-affirm	1- 4	2.25	1.10	.06	-.02	.02	-.01	-.06	-.01	.68**	.72**	-.06	.42**	.57**	-.06	.40**	
15	Thrt X SIC Manip X Self-affirm	1- 8	3.42	2.02	.08	.02	.06	-.07	.51**	-.01	.54**	.56**	.36**	.77**	.46**	.42**	.80**	.79**

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

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