

**A Mediated-Moderation Model of Stereotype Threat and Performance Pressure Effects**

by

Jimmy Louis Johnson

A dissertation submitted to the Graduate Faculty of  
Auburn University  
in partial fulfillment of the  
requirements for the Degree of  
Doctor of Philosophy

Auburn, Alabama  
August 4, 2012

Copyright 2012 by Jimmy Louis Johnson

Approved by

Daniel Svyantek, Chair, Professor of Psychology  
Malissa Clark, Assistant Professor of Psychology  
Jinyan Fan, Assistant Professor of Psychology  
Ana Franco-Watkins, Assistant Professor of Psychology

## Abstract

The present study examined the effectiveness of a model of stereotype threat and performance pressure effects. Specifically, it was proposed that stereotype threat and performance pressure are related to math test performance, and that this relationship was moderated by competitiveness. Additionally, it was proposed that the interactions among these three predictor variables were mediated by test-related perception variables (e.g., test-taking motivation and anxiety). Female participants were assigned to one of four conditions (stereotype threat, performance pressure, combination, control) and given a math ability test. Information concerning their individual levels of competitiveness and perceptions of tests was also collected. Results indicated that only performance pressure was significantly related to math test performance, making it the stronger predictor when compared to stereotype threat. Additionally, perceptions of stereotype threat and performance pressure both had an indirect effect on test performance through the proposed mediator variables, but the effect was in the opposite direction to what was expected. These findings suggest that although it may be difficult to consistently manipulate feelings of threat and pressure, those feelings may actually provide a performance advantage rather than a hindrance.

## Table of Contents

Abstract.....	ii
List of Tables .....	iv
List of Figures.....	v
Introduction.....	1
Origins of Stereotype Threat.....	3
Stereotype Threat as an Explanation.....	5
Susceptibility to Stereotype Threat.....	6
Consequences of Stereotype Threat.....	8
Activating Stereotype Threat .....	9
Performance Pressure.....	13
Susceptibility to Performance Pressure .....	14
Activating Performance Pressure.....	15
Moderators .....	17
Competitiveness.....	17
Mediators .....	20
Mediated-Moderation .....	22
Method .....	23
Results.....	26
Discussion.....	31

References.....	38
Appendix A: Means, Standard Deviations, Intercorrelations and Models .....	47
Appendix B: Math Ability Test and Post-Test Questionnaire .....	69

## List of Tables

Table 1: Means, Standard Deviations, Intercorrelations for all Test Variables.....	48
Table 2: Factorial ANOVA: Stereotype Threat and Performance.....	49
Table 3: Multiple Regression: Perceived Stereotype and Perceived Performance Pressure .....	50
Table 4: Hypothesis 2 Excluded Variables.....	51
Table 5: Indirect Effects of Condition .....	52
Table 6: Indirect Effects of Perceived Stereotype Threat and Perceived Performance Pressure ..	53
Table 7: Indirect Effects of Interaction Variables.....	54

## List of Figures

Figure 1: Conceptual Model of Stereotype Threat Effects .....	55
Figure 2: Conceptual Model of Moderation: Stereotype Threat and Competitiveness .....	56
Figure 3: Conceptual Model of Moderation: Performance Pressure and Competitiveness.....	57
Figure 4: Conceptual Model of Indirect Effect: Stereotype Threat .....	58
Figure 5: Conceptual Model of Indirect Effect: Performance Pressure.....	59
Figure 6: Statistical Model of Mediated Moderation:.....	60
Figure 7: Indirect Effect of Condition .....	61
Figure 8: Indirect Effect of Stereotype Threat and Performance Pressure .....	62
Figure 9: Indirect Effect of Perceived Stereotype Threat * Competitiveness .....	63
Figure 10: Indirect Effect of Perceived Performance Pressure * Competitiveness .....	64
Figure 11: Indirect Effect of Stereotype Threat Condition * Competitiveness .....	65
Figure 12: Indirect Effect of Performance Pressure Condition * Competitiveness.....	66
Figure 13: Indirect Effect of Combination Condition * Competitiveness.....	67
Figure 14: Average Performance by Racial Subgroup Across Conditions.....	68

## Introduction

In recent decades, research has demonstrated that the prevalence of negative stereotypes has potentially damaging effects for individuals who are targeted by those stereotypes. Steele and Aronson (1995) were among the first to demonstrate this effect in their study that examined racial subgroup differences in scores on tests of cognitive ability. In their study, White participants scored significantly higher than their Black counterparts. They concluded that the observed differences were the result of *stereotype threat*, which occurs when the presence of a negative stereotype has an adverse effect on a targeted individual's performance in the stereotyped domain. Additionally, the effects of this threat are typically thought to be independent of the targeted individual's true ability level in the domain in question. The effects of stereotype threat have been demonstrated using a number of different stereotypes such as those that describe ethnic minorities as inferior to Whites in terms of intellectual abilities and those that describe women as not as good as men at math (Steele, Spencer, & Aronson, 2002). The message that the theory of stereotype threat conveys is powerful: members of stigmatized social groups may be constantly at risk of underperformance in high-stakes testing situations (Steele, 1997). However, individuals need not always be *stigmatized* to fall victim to stereotype threat. Any group of individuals who can be unfavorably compared to a salient outgroup may be susceptible. For these reasons, stereotype threat is a very important concern in the field of psychology, as well as at the societal level. It is therefore important that we identify variables that may lessen the damaging effects of stereotype threat.

Interestingly, stereotype threat may not be the only explanation for subgroup differences in test performance. Another variable, *choking under pressure* (e.g., Baumeister, 1984; Beilock & Carr, 2001) shares many similarities with stereotype threat and can often result in similar outcomes with respect to task performance. Baumeister (1984) defines choking under pressure as a decrease in performance on a task in situations where successful performance is important to the individual performing the task. Put another way, when individuals are in a high-stakes situation with a great deal of performance pressure, they often fail despite having the ability to perform successfully. Using this logic, one might expect to see more failures during a live concert than during the dress rehearsal the previous evening because the pressure to perform is greater in front of a live audience. Stereotype threat and performance pressure have both been proposed as explanations for poor performance in stressful situations (Beilock, 2008), but they have not yet been investigated simultaneously to determine which is the stronger predictor of task performance, or the more appropriate explanation for the achievement gap.

Although previous research has demonstrated the viability a number of moderators (e.g., domain identification, test diagnosticity), to date, no one has investigated whether competitiveness moderates the relationship between stereotype threat and performance. The proposed moderator is important because it has the potential to provide utility for *any* negative stereotype and targeted group. Additionally, because stereotype threat and performance pressure have been shown to result in similar performance-based outcomes, competitiveness will also be investigated as a moderator of the relationship between performance pressure and task performance. The stereotype chosen for this study proposes that women are inferior to men in terms of math intelligence.



The primary objectives of the present study are as follows: (a) to compare the effects of stereotype threat and performance pressure directly in order to determine which is the best predictor of performance, (b) to test the moderating effect of competitiveness on the relationship between pressure condition (e.g., stereotype threat or performance pressure) and task performance and (c) to investigate test-related perception variables (e.g., test-taking anxiety, test-taking motivation) as mediators of the interactive effect of pressure condition and competitiveness on task performance. Figure 1 depicts the conceptual model to be tested in the current study.

### **Origins of Stereotype Threat**

Stereotype threat was first introduced by Steele and Aronson (1995) to explain the *achievement gap* between Black and White students. The achievement gap manifested itself as a 1.0 standard deviation discrepancy in average scores on standardized tests of cognitive ability between Blacks and Whites. Up until this point, many of the explanations for the achievement gap were inflammatory, claiming that Blacks are cognitively inferior to Whites and that these differences are very resistant to change (e.g., Gottfredson, 1988; Herrnstein & Murray, 1994). Stereotype threat provided an alternative explanation that focused on test perceptions rather than differences in innate ability levels. Consequently, stereotype threat has been proposed as an explanation for a number of subgroup differences. One of the most common subgroup differences addressed by stereotype threat aside from racial differences in cognitive ability is gender difference in math ability.

While arguably not as salient and widespread as racial stereotypes focusing on cognitive ability, a number of gender stereotypes exist, including the stereotype that designates females as inferior to men in terms of math ability. Girls report greater self-competence in verbal domains,

whereas boys report greater self-competence in and valuing of mathematics and science (Bornholt, Goodnow, & Cooney, 1994; Eccles, Wigfield, Harold, & Blumenfeld, 1993; Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002; Marsh et al., 2005; Marsh & Yeung, 1997; Stetsenko et al., 2000; Wigfield et al., 1997). It is therefore logical that on the average, males should perform better on tests of math ability than females. The consequences of this phenomenon may not be as apparent because math is only one facet of intelligence. Furthermore, students generally select college majors and subsequently careers that correspond to subjects with which they feel the most competent. For (female) students who are not comfortable with math, college majors may be chosen from paths that do not necessitate an extensive math curriculum such as English, Education, or Law. Accordingly, women are typically underrepresented in math-heavy careers such as engineering. However, a number of college entrance exams include a math subsection. The Scholastic Aptitude Test (SAT), American College Test (ACT), and Graduate Record Exam (GRE) all include sections that focus on math ability. Because these tests are frequently used as selection criteria for college admissions and scholarships, math ability is both important and consequential.

There are a number of explanations for the discrepancy in math ability scores between males versus females. Many theoretical models that focus on this gender difference begin with the assumption that females underperform compared to males and attempt to explain the causes of this phenomenon. For instance, Eccles (1987) proposed an Expectancy x Value model to explain the phenomenon that males and females choose different math courses in high school, with males typically choosing math classes that are more advanced and challenging. Similarly, Fennema and Peterson (1985) constructed an autonomous learning behavior model which suggests that failure to participate in independent learning in mathematics contributes to gender

differences in math performance. Additionally, they proposed that teachers exert more control of female students than male students and also encourage independent thinking more frequently for male students than female students. This pattern of treatment and minimal independent thinking also predicts feelings of incompetence in females, presumably in academic areas such as mathematics, as compared to their male counterparts. Therefore, Fennema and Peterson's (1985) model contains both behavioral elements, and environmental elements that shape those behaviors. Finally, similar to the more inflammatory explanations of the achievement gap in cognitive ability between African American students and White students (e.g., Gottfredson, 1988; Hernstein & Murray, 1994), a number of researchers propose biological differences as a possible explanation. For example, Halpern (1986) suggests that gender differences in lateralization of brain function may be yet another explanation.

### **Stereotype Threat as an Explanation**

In contrast to the above explanations, stereotype threat theory proposes that the presence of a (negative) stereotype can have an adverse effect on a targeted individual's performance in the stereotyped domain (e.g., Steele and Aronson, 1995; 1998). The fact that stereotypes are the topic of interest also implies the presence of a group that is targeted by the stereotype, as well as an outgroup to which the stereotyped group is compared. Consequently, the foundation for stereotype threat comes primarily from social identity theory (Tajfel & Turner, 1979; 1986). This theory proposes that individuals identify themselves and others using salient characteristics (e.g., socioeconomic status, religion, political affiliation, race), and subsequently surround themselves with similar and like-minded individuals while staying away from those whose beliefs and behavioral tendencies are different from their own. As a result, group members are provided with reinforcement for their beliefs and characteristics and their group identity is solidified. This same

reinforcement also helps group members to distinguish their ingroup from dissimilar individuals, known as the outgroup. Group members tend to desire equal if not favorable comparisons to their outgroups. Unfavorable comparisons can be very threatening, which may cause heightened group tension and even overt backlash (Rudman & Fairchild, 2004).

The threatening unfavorable comparisons provide the necessary fuel for stereotype threat to occur. In the case of the present study, the stereotype in question describes females as inferior to males in terms of math ability. In testing situations, the threat of confirming this stereotype is often enough to adversely affect performance and result in the very differences in math performance that the stereotype predicts. Put another way, the negative stereotype in question is self-perpetuating, to the extent that its very existence results in disparate group performance that in turn prolongs its existence.

### **Susceptibility to Stereotype Threat**

Although stereotype threat can affect any person who is targeted by a stereotype, there are a number of factors that make targets more susceptible. For instance, stereotypes are arguably most threatening to targeted individuals who endorse them as opposed to those who disagree with them. As an example, Black students who accept the information in damaging stereotypes as valid (e.g., Blacks are inferior to Whites with respect to cognitive ability) are likely to be more susceptible and less resistant to the negative effects of stereotype compared to Black students who view the stereotype as inaccurate and consequently attempt to disprove it. In the case of the stereotype concerning the math abilities of women, recent research investigated the extent to which women actually endorse these beliefs. Blanton, Christie, and Dye (2002) found that although the tendency towards stereotype endorsement was not prevalent in their sample of women, 41% of their sample stated that there is at least “some truth” to the stereotype that men

are superior to women in terms of math ability. This finding is somewhat surprising because social identity theory predicts that women should be more likely to endorse stereotypes that compare themselves in a favorable light with respect to men (Tajfel & Turner, 1979; 1986). Additionally, Nosek, Banaji, and Greenwald (2002) found that in their sample of college women, there was a strong implicit association between the words *math* and *maleness*. Interestingly, the sample chosen for this study included females majoring in math-related fields, who presumably identified strongly with that domain. The strength of this association was correlated with poorer performance on the quantitative section of the SAT.

Beyond stereotype endorsement, domain identification is an additional factor that is important to assess when predicting how susceptible a targeted individual will be to the effects of stereotype threat. Individuals who strongly identify with the targeted domain of performance or expertise are susceptible to the threat of confirming a negative stereotype because the strength of stereotype threat effects depends upon the extent to which one's self-worth is tied to the outcomes of the individual's performance in the stereotyped domain (Steele et al., 2002). As an example, individuals who consider themselves to be intelligent will be susceptible to negative stereotypes that predict unfavorable group differences in cognitive ability between their ingroups and certain outgroups. Contrastingly, individuals who do not identify strongly with the domain in question will likely not be affected by stereotype threat because failure in the domain will have little effect on that person's self-worth. Steele and colleagues (2002) demonstrated that individuals who *moderately* identify with the stereotyped domain experience the highest degree of stereotype threat, while individuals who report a *low* level of domain identification experience the lowest degree of stereotype threat and individuals who *highly* identify with the stereotyped domain experienced a medium degree of threat.

Task difficulty is yet another variable that must be considered when determining the likelihood that an individual will fall victim to the negative effects of stereotype threat. Stereotype threat theory predicts that targeted group members are particularly likely to experience stereotype threat when a test or task in the stereotyped domain of performance is challenging as opposed to easy (e.g., Steele & Aronson, 1995; 1998). Presumably, difficult and challenging tasks will require participants to devote a greater amount of cognitive resources than simple tasks. Consequently, the chance for cognitive interference from a threatening stereotype is more likely because (a) the threatened participant will not have enough cognitive resources to focus on task performance and worry about the consequences of poor performance simultaneously, and (b) the resources that are used to worry will not be available to devote to task performance (Steele & Aronson, 1995; Steele et al., 2002). Previous research has demonstrated that in samples of both female and Black participants, highly difficult tasks produce higher levels of stereotype threat and subsequently lower levels of performance than moderately difficult tasks (Nguyen & Ryan, 2008).

In sum, the aforementioned research demonstrates that stereotype threat does not affect all people equally. Individuals who believe that the stereotype that targets them is valid, as well as individuals who moderately identify with the stereotyped domain are likely to experience a heightened degree of stereotype threat. Also, stereotype threat tends to be higher in situations where the targeted task is difficult and challenging.

### **Consequences of Stereotype Threat**

In addition to test performance, stereotype threat has been demonstrated to result in a number of negative consequences for stereotyped individuals. With African American employees, stereotype threat can result in lowered feedback seeking and higher discounting of

feedback (Roberson, Deitch, Brief, & Block, 2007). This finding is especially consequential because it could lead to lowered job performance in the long run as a result of a refusal to integrate constructive feedback into inefficient or unsafe work behaviors. With women, math anxiety caused by negative stereotypes correlates negatively with interest in scientific careers (Chipman, Krantz, & Silver, 1992). Additionally, acceptance of math stereotypes in women may cause them to choose not to pursue math-related careers (Kiefer & Sekaquaptewa, 2007). Taken together, these findings indicate that as women become more aware of negative stereotypes that target them, they may be less likely to pursue careers in fields such as biology, medicine, engineering, or physics. In sum, stereotype threat is an issue that affects not only immediate performance, but also affects perceptions outside of the performance situation.

### **Activating Stereotype Threat**

Although stereotype threat presumes that the stereotype in question is widespread enough to be threatening, there are certain situations that are believed to “activate” the threat. Steele and Aronson (1995; 1998) found that presenting a difficult test as “diagnostic of ability” produced enough threat in academically successful Blacks at Stanford University to disrupt performance on an ability test. However, when the test was presented as nondiagnostic or inconsequential (e.g., no mention of cognitive ability) the threat was absent and there was no observed discrepancy between the scores of Black and White students. This effect was also replicated for female test-takers on difficult math ability tests (e.g., Davies, Spencer, Quinn, & Gerhardstein, 2002; Schmader & Johns, 2003; Spencer, Steele, & Quinn, 1999). In practice, these effects are sometimes difficult to transfer to real-world settings. Difficult tests of any variety usually imply some sort of consequence or indication of ability. Furthermore, falsely telling test-takers that an important test is inconsequential is unethical.

Nguyen and Ryan (2008) categorize stereotype-threat activating cues based upon how explicitly the cue is perceived. Cues can be described as blatant, moderately explicit, or indirect and subtle. As expected, blatant cues are the most explicit cues. The negative information of the stereotype is clearly and explicitly conveyed to targeted individuals prior to their performance in the stereotyped domain (e.g., math test). In other words, targeted individuals are made aware of the negative stereotype through a conscious process (Nguyen & Ryan, 2008). As an example, stating that Whites tend to perform better than Blacks or that men tend to perform better than women would be considered blatant cues (e.g., Aronson, Lusting, Good, Keough, Steele, & Brown, 1999; Cadinu, Frigerio, Impagliazzo, & Latinotti, 2003). Also, administering a stereotype threat questionnaire or giving information favoring the outgroup *before* performance would be considered a blatant cue (e.g., Bailey, 2004).

In the case of moderately explicit cues, the message of a negative stereotype is given within the instructions of an ability test. However, the instructions merely state that there is a *group difference*. The direction of the group difference is purposely left open to the interpretation of the test-takers. This type of cue is still considered to be explicit because the negative stereotype becomes salient to targeted individuals through a conscious mechanism (Nguyen & Ryan, 2008). For instance, stating that generally men and women perform differently on math ability tests or that certain groups of people perform better than others on cognitive ability tests are examples of moderately explicit cues (e.g., Edwards, 2004; Pellegrini, 2005).

Finally, indirect and subtle cues are the least explicit of the three classifications of threat-activating cues. With indirect and subtle cues, the message of group differences in the stereotyped domain is not explicitly stated. Instead, the context of tests, test takers' subgroup memberships, or test taking experience is manipulated. As opposed to the other two



classifications, indirect and subtle cues result in the negative stereotype becoming salient to test takers through an automatic or subconscious process (Nguyen & Ryan, 2008). Race/gender priming is an example of an indirect cue that can be accomplished by giving a pretest questionnaire that contains racial information (e.g., Oswald & Harvey, 2000-2001; Schmader & Johns, 2003). Additionally, emphasizing test diagnosticity or stressing the evaluative nature of the test is considered an indirect cue (e.g., Ployhart, Ziegart, & McFarland, 2003).

Research has demonstrated that each of the aforementioned cues may have different effects on task performance in threat-loaded situations. Researchers suggest that moderately explicit cues typically have a more detrimental effect on task performance in the stereotyped domain than explicit cues (e.g., Bargh, 1997). Explicit cues can result in a phenomenon known as *stereotype reactance*, which is essentially the opposite effect of stereotype threat (Oswald & Harvey, 2000-2001). When a negative stereotype is made explicit, it may be perceived as a limit to the stereotyped individual's ability and elicit compensatory behaviors that are inconsistent with the stereotype (Kray, Thompson, & Galinsky, 2001). As a result, targeted individuals perform better than expected in the presence of a threatening situation. That being said, stereotype threat manipulations in laboratory settings should be conducted using moderately explicit threat-activating cues.

Stereotype threat can also be activated using cues that are so subtle that they may often be ignored by those who are unfamiliar with stereotype threat. Researchers warn that in many cases, the word "test" alone is sufficient enough to elicit threat (Steele & Davies, 2003). Additionally, mentioning race or gender may remind test-takers of negative stereotypes that target those groups, which has led to researchers such as Stricker and Ward (1998) to collect demographic

information (e.g., ethnicity, gender, age) *after* the test. In short, stereotype threat is a very volatile construct.

In addition to categorizing stereotype threat-activating cues into three classifications, Nguyen and Ryan (2008) also identified a number of stereotype threat-removal strategies and designated each of them as *explicit* or *subtle* strategies. As expected, explicit strategies involve an overt and obvious mention of information designed to remove threat. For instance, giving a handout prior to performance that favors the targeted group or educating participants about the stereotype threat phenomenon are both examples of explicit strategies (e.g., Bailey, 2004; Guajardo, 2005). Contrastingly, subtle strategies do not involve explicit statements of stereotypical or counterstereotypical information, and can include techniques such as describing a threatening test as a problem-solving task or showing television commercials with targeted group members in astereotypical roles prior to task performance (e.g., Steele & Aronson, 1995; Davies et al., 2002). Research suggests that of these two alternatives, explicit stereotype threat-removal strategies should eliminate the negative effects of stereotype threat more efficiently than implicit strategies because explicit strategies are more likely to motivate individuals to exert more effort in an attempt to avoid being stereotyped (Spencer, Fein, Strahan, & Zanna, 2005). However, this information is unlikely to be used very often in practical or applied settings, because hiring organizations typically do not utilize explicit threat-removal strategies (Nguyen & Ryan, 2008).

With regard to manipulating stereotype threat effects in experimental settings, it is important to make decisions based upon the stereotype in question as well as the performance domain. In the seminal study by Steele and Aronson (1995), participants in the stereotype threat condition (diagnostic) were informed that a set of questions from the verbal section of the GRE

would provide a “genuine test of verbal abilities and limitations.” Further, participants were told that the study itself was concerned with “various personal factors involved in performance on problems requiring reading and verbal reasoning abilities.” However, in the non-diagnostic condition, test-takers were told that test items were being used to familiarize them with items that they may encounter on future tests. Additionally, they were informed that the purpose of the study was “to better understand psychological factors involved in verbal problem solving.” Clearly, the mention of *ability* was the primary characteristic that distinguished the stereotype threat conditions from another. When a test is described as diagnostic of intellectual (verbal) ability, negative stereotypes concerning intellectual ability become relevant to targeted individuals and the threat of confirming that stereotype through poor performance is then realized.

The preceding example of research design is relevant for both historical and practical reasons. However, it focuses on Black/White subgroup differences, and the stereotype that predicts those differences. Consequently for the purposes of the present study, elements of that research designs must be altered in order to test gender differences in math ability. Studies by Schmader (2002) and Perry and Skitka (2009) provide the basis for the research design that will be used in the present study. Although the relationship between stereotype threat and performance has been previously tested in numerous studies, it is necessary to test this hypothesis before proceeding with subsequent hypotheses which serve as the basis for the primary objectives. Accordingly, I propose the following:

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

## **Performance Pressure**

Performance pressure is a related variable that has resulted in outcomes that are very similar to stereotype threat. The desire to perform as well as possible in a situation with a high degree of personally-felt importance creates *performance pressure* (Baumeister, 1984; Hardy, Mullen, & Jones, 1996). Performers strive to do their best in situations with a high degree of performance pressure, yet often they *choke*, and perform at a lower level than would have otherwise been expected (Beilock & Carr, 2001). This effect can be seen across many domains including sports (e.g., missing a wide open game-winning layup), music (e.g., forgetting the words to the National Anthem before a televised sporting event), and academics (e.g., confusing Austria with Australia on a geography test). Although performance pressure is studied more as a cognitive variable and stereotype threat is studied more as a social variable, both stereotype threat and performance pressure have been used to explain poor performance in high-stakes situations.

### **Susceptibility to Performance Pressure**

While the effects of stereotype threat are specific to only the targets of the stereotypes in question, performance pressure affects individuals more consistently across racial and gender groups because susceptibility is contingent upon working memory capacity. It has been demonstrated that individuals who are high in working memory capacity experience more a detriment in performance in pressure-laden situations than low capacity individuals (Beilock, 2008). Working memory is important because when performance pressure is high, affected individuals are hindered by anxiety and intrusive thoughts that tie up working memory resources that could ordinarily be used for successful performance (Beilock, 2008; DeCaro, Rotar, Kendra, & Beilock, 2010). This effect is similar to intrusive thoughts that Steele and Aronson (1995) identified under conditions of stereotype threat in which targeted performers worry not only

about the difficulty of the task, but also about the possibility that the task is only difficult for targeted group members.

### **Activating Performance Pressure**

Much like stereotype threat, the amount of performance pressure that an individual experiences depends on a number fairly standard environmental factors. The first of these factors is social evaluation. When an individual is made aware that his or her performance will be observed and critiqued by experts in the performance domain, performance pressure is likely to increase. This social evaluation can take the form of a committee of professors listening to a research proposal from an undergraduate studying in their major, or high-school aged pitcher playing in the state championship while several college and professional scouts are in attendance. In both cases, performers are concerned not only with their performance, but also with the how good the performance is appears to others.

The next factor that influences the amount of performance pressure that an individual will feel is monetary compensation. While providing a monetary reward that is contingent upon exceptional performance would seem to increase task motivation, the fear of missing out on that compensation elicits a great deal of pressure during the performance. As an example, a friendly game of poker may elicit a greater deal of performance pressure for the players if they each have to buy into the game for \$20, than if their chips are free and have no associated monetary value. When rewards such as money are at stake, performers generally feel a large amount of pressure to perform successfully, and thus may find themselves worrying about the monetary consequences of losing or simply not winning.

The final factor that influences the amount of performance pressure that an individual will feel is peer pressure. In this case, the task in question is typically team-based and

interdependent: requiring optimal performance from all team members to be successful. Consequently, individual team members feel a great deal of pressure to perform and let the team down. Failure on the part of one individual can lead to failure for the entire team, despite the possibility that other team members may have performed successfully. An example of this effect could be seen in a doubles tennis match in which both team members must perform exceptionally to ensure a victory. This same effect is less likely to be seen in a 100 person choir in which many individuals are responsible for the (singing) task, and the absence of one person could easily go unnoticed. When another individual or a team of other individuals is depending on the performance of one individual for the joint success of a single task, that one individual should be expected to experience a great deal of pressure to perform.

As stated previously, stereotype threat and performance pressure are two variables that share a number of important similarities, and that have generally resulted in similar performance-based outcomes. Although they have been discussed in tandem (e.g., Beilock, 2008), they have not yet been investigated simultaneously within the same study. Perhaps it is the case that stereotype threat is a type of performance pressure in which the incentive to perform is not social evaluation, money, or peer pressure, but to disconfirm a negative stereotype about performer's group. This would make it difficult to predict which of the variables would have a more negative impact on performance, but the presence of one or both variables should result in lower performance relative to the absence of both variables. Accordingly, I propose the following:

**Hypothesis 1b:** Test performance will be higher in the control (no-stereotype threat) condition than in the performance pressure condition.

**Hypothesis 1c:** Test performance will be higher in the control (no-stereotype threat) condition than in the combination condition (stereotype threat + performance pressure).

## **Moderators**

Although stereotype threat effects have been demonstrated and replicated by a number of researchers (e.g., Steele & Aronson, 1995; Davies et al., 2002; Schmader & Johns, 2003). A number of researchers have not found support for the theory (e.g., Oswald & Harvey, 2000-2001; Stricker & Ward, 2004). The mixed findings suggest the presence of moderating effects for stereotype threat (Nguyen & Ryan, 2008). Moderator variables are important primarily because they may provide buffer conditions for the negative effects of stereotype threat. In other words, identifying and manipulating moderators can make individuals more or less susceptible to stereotype threat. Furthermore, as mentioned earlier, explicit stereotype threat removal strategies are usually either unethical or not feasible to implement given the circumstances of the performance domain. There are likely few organizations that would take the time to educate all potential employment candidates about stereotype threat, and even fewer that would overtly mislead candidates about the importance of the selection criteria (e.g., intelligence test) which may elicit subgroup differences in performance as a result of stereotype threat. Therefore, the necessity for identifying variables that may lessen the effects of stereotype threat is paramount, particularly for applied settings.

A number of moderator variables have been identified in previous research including stigma consciousness (Edwards & Arthur, 2004), racial group identity (Davis, Aronson, & Salinas, 2006), defensive pessimism (Perry & Skitka, 2009), and coping sense of humor (Ford, Ferguson, Brooks, & Hagadone, 2004). In addition to these variables, competitiveness is another potential moderator variable that may provide utility to those who may fall victim to stereotype threat.

## **Competitiveness**

Competitiveness is an important concept that is recognized as playing a significant role in interpersonal processes (Gough, 1987). Although research has yielded a number of different definitions, competitiveness can be generally described as a desire to win or to do better than others in interpersonal situations (Helmreich & Spence, 1978). With respect to stereotype threat, competitiveness is a character trait that should shield individuals from the negative effects of threatening situations for a number of reasons. Individuals who are high in competitiveness should be more likely to reappraise a threat-loaded situation as a challenge. The *challenge* designation is similar to the manipulation used by Steele and Aronson (1995) and referenced previously in the discussion of sense of humor as a moderator variable. Their study showed that the manipulation effectively reduced the experience of stereotype threat for targeted individuals. Additionally, instead of being concerned with the threat of *confirming* a negative stereotype, targeted individuals who are competitive may be more concerned with the self-esteem boost that would result from *disconfirming* the negative stereotype. In fact, competitive individuals who are targeted by a negative stereotype may thrive in the so-called threatening situation that other less competitive-minded individuals find to be so debilitating. When confronted with implications that they are inferior to an outgroup in terms of domain performance (e.g., intelligence test, math ability test), competitive individuals' desires to win and to do better than the outgroup may be fueled by the stereotype instead of hindered by it. Further, the outgroup provides a convenient "target" at which stereotyped individuals can focus their efforts.

In addition to not feeling threatened by a negative stereotype, competitive individuals may also report higher levels of task motivation than non-competitive individuals when confronting a threat-loaded situation. In this case, however, targeted individuals are focused not only on succeeding (e.g., demonstrating expertise in the performance domain), but also with



proving that their ingroup is comparable or superior to the outgroup of comparison, and thus disconfirming the negative stereotype. Social identity theory provides further support for this notion because it predicts that individuals want equal or favorable comparisons of their ingroup to outgroups (Tajfel & Turner, 1979; 1986). Accordingly, competitive individuals should be more likely to work harder to demonstrate the comparable or superior status of their ingroups in favor of utilizing other threat-reducing strategies such as recategorizing their ingroup membership. This heightened level of motivation should manifest itself as better performance (e.g., higher test scores) than non-competitive individuals when stereotype threat is present.

Additionally, because stereotype threat and performance pressure have been demonstrated to result in similar outcomes, it is not unreasonable to suspect that competitiveness may display similar moderating effects on the relationship between performance pressure and test performance. That is, individuals who enjoy competition are likely to be competitive, and competitions typically elicit some degree of pressure to perform well. Competitive individuals may therefore be more likely to thrive in pressure-laden situations than non-competitive individuals. Accordingly, I propose the following:

**Hypothesis 2a:** Competitiveness will moderate the relationship between stereotype threat and performance such that under conditions of stereotype threat, test performance will be higher when individuals report high levels of competitiveness as opposed to low levels of competitiveness.

**Hypothesis 2b:** Competitiveness will moderate the relationship between performance pressure and performance such that under conditions of performance pressure, test performance will be higher when individuals report high levels of competitiveness as opposed to low levels of competitiveness.

## **Mediators**

In the past, researchers have attempted to identify mediator variables of the stereotype threat-performance relationship to shed light on the specific mechanism that results in lowered performance. Steele and Aronson (1995) explained that anxiety is the primary underlying mechanism for the effects of stereotype threat. In addition to the anxiety associated with taking a high-stakes test of ability, the threat of confirming a negative stereotype through poor performance elicits additional anxiety that interferes with normal mental processing during the testing situation. While taking a difficult test, targeted individuals focus on the self-significance of their inability to answer questions rather than attempting to identify correct answer. That is, when targeted female students encounter difficult questions that they are unable to answer, they are less concerned with whether the question is difficult and more concerned with whether the question is difficult because they are female. Consequently, targeted students spend more time on the task, yet end up answering fewer questions than non-targeted students. (Steele & Aronson, 1995).

The influence of examinee perceptions on test performance has been increasingly studied in the last several years (Ryan & Ployhart, 2000) and driven by findings that various perceptions (e.g., face validity) are related to real outcomes such as test performance or decision to withdraw from the job application process. Some researchers have posited that group differences in standardized test performance can be attributed to differences in perceived face validity, fairness, predictive validity, belief in the utility of tests, and self-efficacy (e.g., Chan & Schmitt, 1997; Chan, Schmitt, DeShon, Clause, & Delbridge, 1997; Edwards & Arthur, 2007; Gilliland, 1994; Ryan, 2001). It may be that stereotype threat decreases test performance for targeted individuals because the threatening condition elicits negative perceptions related to the test and/or testing

situation. Ployhart et al. (2003) obtained evidence that test-taking perceptions (e.g., anxiety, face validity, and motivation) were significant mediators of the race relationship between stereotype threat and performance.

The studies mentioned above all share one important similarity: they all focused on racial subgroup differences in cognitive ability, and the related stereotype that predicts that Blacks are cognitively inferior to Whites. Consequently, a number of the test-related perception variables (e.g., perceived fairness) reflect decades of racial tension. With respect to gender differences in math ability, these mediator variables should not be relevant because the interpersonal dynamic between racial subgroups is fundamentally different than the interpersonal dynamic between genders. For instance, racial subgroup differences have been attributed to unfair education advantages granted to White children with respect to the types of schools that children from each group attend (Berry, 2003). This difference would not apply to gender differences because there are both males and females in both of those educational groups. However, anxiety and motivation are two task-related perception mediator variables that should apply regardless of the targeted group or the stereotype in question. Also as previously mentioned, if it is in fact the case that stereotype threat and performance pressure are essentially the same thing, then it is reasonable to assume that the same task-related perception variables that mediate the relationship between stereotype threat and performance should also mediate the relationship between performance pressure and performance. Accordingly, I propose the following:

**Hypothesis 3a:** The relationship between stereotype threat and test performance will be partially mediated by test-taking anxiety.

**Hypothesis 3b:** The relationship between stereotype threat and test performance will be partially mediated by test-taking motivation.

**Hypothesis 3c:** The relationship between performance pressure and test performance will be partially mediated by test-taking anxiety.

**Hypothesis 3d:** The relationship between performance pressure and test performance will be partially mediated by test-taking motivation.

### **Mediated-Moderation**

Mediated-moderation is a somewhat murky process that has been introduced in recent years (e.g., James & Brett, 1984; Baron & Kenny, 1986) that combines mediation and moderation into one procedure. In short, mediated moderation occurs when a moderator affects the magnitude of the initial predictor variable on the mediator variable (Muller, Judd, & Yzerbyt, 2008). Put another way, mediated-moderation is present when the interactive effect of X1 and X2 results in a change in the mediating variable, which in turn leads to a change in Y. In the case of the present study, I have proposed that competitive individuals will not experience stereotype threat or performance pressure as much as non-competitive individuals. The lowered experience of those two variables should result in low anxiety and high motivation, which should in turn predict task performance. Therefore, I propose the following:

**Hypothesis 4a:** The interactive effects of stereotype threat and competitiveness on test performance will be partially mediated by test-taking anxiety.

**Hypothesis 4b:** The interactive effects of performance pressure and competitiveness on test performance will be partially mediated by test-taking anxiety.

**Hypothesis 4c:** The interactive effects of stereotype threat and competitiveness on test performance will be partially mediated by test-taking motivation.

**Hypothesis 4d:** The interactive effects of performance pressure and competitiveness on test performance will be partially mediated by test-taking motivation.

## **Method**

### **Participants**

The sample consisted of 153 female participants at a large southeastern university. Respondents ranged in age from 18 to 28 years. Of the participants, 2 were American Indian, 2 were Asian, 18 were African American, 5 were Hispanic, 118 were White, and 3 reported “Other” for their race/ethnicity. There were 32 freshmen, 26 sophomores, 50 juniors, 39 seniors, 2 graduate students, and 4 participants who did not report classification. A total of 5 participants were removed from the sample because of incomplete data.

### **Procedure**

Participants were randomly assigned to one of four conditions: stereotype threat, performance pressure, combination (stereotype threat + performance pressure) and control. The conditions were distinguished by the specific instructions that the participants receive. In the stereotype threat condition, participants were told that the test is “indicative of your mathematic ability.” Further, participants were told that “in the past, math tests such as the one you are taking have produced gender differences in math ability” and “because we are comparing women’s scores to men’s, each of your scores will also be used as an indicator of women’s or men’s math ability in general.” Finally, participants were required to indicate their gender before beginning the test. This design was adapted from research designs used by Schmader (2002) and Perry and Skitka (2009). The manipulation used moderately explicit cues because they have been demonstrated to have more damaging effects than blatant or indirect cues when used to elicit stereotype threat (Nguyen & Ryan, 2008). The instructions telling the participants that “math tests such as the one you are taking have produced gender differences in math ability” can be considered to be moderately explicit cue because there is no mention of the hypothesized

direction of these gender differences. Additionally, instructions telling the participants that the test is indicative of math ability and requiring participants to indicate their gender before taking the test are both examples of subtle cues because there is no mention of a possible comparison between male and female scores. Blatant cues (e.g., men perform better than women on math tests) were not included because of the possibility of a stereotype reactance effect (e.g., Oswald & Harvey, 2000-2001).

In the performance pressure condition, participants were told that the questions are evaluative of *problem-solving ability* instead of math ability. In order to provide a monetary incentive, participants were told that if they scored at least an 80% on the task, they would receive a \$15 reward. To elicit peer pressure, the participants were told the following:

*Attaining the monetary award is a team effort, and you have been randomly paired with another participant who has already completed the study with a score of at least 80%. In order for both of you to receive the award, you must also score an 80%. If you fail to do so, neither you nor your randomly assigned partner will receive \$15. You will be notified at the conclusion of the study if your score qualifies you and your partner for the reward.* This design was adapted from a previous research design used by DeCaro et al. (2010).

In the combination condition, participants received instructions from both the stereotype threat and performance pressure conditions. Finally, in the control condition, participants were told that “the questions are designed to investigate the strategies that they use to solve world problems” and that “although no feedback will be provided, they should still give their best effort.”

All participants were given the instructions that correspond to their condition, and then given 25 minutes to complete a math ability test. The study employed a 2x2 design with the four

conditions mentioned above. After the math ability test, participants completed the remainder of the measures in the following order: perceived stereotype threat, perceived performance pressure, test-taking motivation, test-taking anxiety, and competitiveness.

## **Measures**

**Perceived stereotype threat.** Several subjective indicators (11 total items) of stereotype threat were included as a manipulation check for the stereotype threat conditions: The Perceived Stereotype Threat Scale (Ziegert, Ployhart, & McFarland, 2002), the Post-test Attitudes Survey (McKay, 1999), and the stereotype threat perception measure developed by Sawyer and Hollis-Sawyer (2003). All ratings were made on a 5–point Likert-type scale (1 = *strongly disagree*; 5 = *strongly agree*) and the mean was used as an indicator of perceived stereotype threat. Items can be found in Appendix B.

**Perceived performance pressure.** Perceived performance pressure was measured using the same procedure as Beilock et al. (2004). The following item was included to serve as a manipulation check for performance pressure: “I felt a great deal of pressure to perform well on this test.” Ratings were made on a 5–point Likert-type scale (1 = *strongly disagree*; 5 = *strongly agree*). The item can be found in Appendix B.

**Test–Taking Motivation.** Test–taking motivation was measured using eight items adopted from the Test Attitude Survey (Arvey et al., 1990; Items from Scale 1) which were reworded to measure general test–taking motivation, and examples include “Doing well on this test is important to me” and “I tried my best on this test”. All ratings were made on a 5–point Likert-type scale (1 = *strongly disagree*; 5 = *strongly agree*) and the mean was used as an indicator of test-taking motivation. Items can be found in Appendix B.

**Test-Taking Anxiety.** Test-taking anxiety was measured using 10 items adapted from the TAS (Arvey et al., 1990; Items from Scale 4) which were reworded to measure general test-taking anxiety. Examples of items are “I am not good at taking tests” and “I usually get very anxious about taking tests”. All ratings were made on a 5-point Likert-type scale (1 = *strongly disagree*; 5 = *strongly agree*) and the mean was used as an indicator of test-taking anxiety. Items can be found in Appendix B.

**Competitiveness.** Competitiveness was measured using 8 items from the Interpersonal Competitiveness subscale of the Competitiveness Questionnaire (Griffin-Pierson, 1990). Examples of items are “It is important to me to perform better than others on a task” and “I would want an A because that means that I am better than other people”. All ratings were made on a 5-point Likert-type scale (1 = *strongly disagree*; 5 = *strongly agree*) and the mean was used as an indicator of competitiveness. Items can be found in Appendix B.

**Math Ability.** Math ability was tested using 14 multiple-choice questions from the quantitative section of the SAT. All of the questions were word problems requiring advanced algebraic calculations. Answers were recorded in a free-response format (e.g., not multiple-choice). Items can be found in Appendix B

## Results

Descriptive statistics and bivariate relationships are presented for all variables in Table 1. Data pertaining to levels of perceived stereotype threat and perceived performance pressure were collected as a manipulation check. Math test score correlated significantly with Perceived Stereotype Threat ( $r = .23, p < .01$ ), Perceived Performance Pressure ( $r = .25, p < .01$ ), Anxiety ( $r = -.38, p < .01$ ), and Motivation ( $r = .21, p < .01$ ). While the latter two test-perception variables were expected to correlate in the observed directions with math test score, the positive



correlations between perceived stereotype threat and test score and perceived performance pressure and test score were particularly surprising because they are inconsistent with the expectation that perceptions of stereotype threat and performance pressure should hinder test performance. With regard to the manipulation check variables, there was a significant difference among conditions in terms of perceived stereotype threat  $F(3, 145) = 2.94, p = .04$ . Individuals in the stereotype threat condition reported a significantly higher degree of perceived stereotype threat ( $M = 3.04, SD = .38$ ) than individuals in the performance pressure condition ( $M = 2.81, SD = .42$ ). Levels of perceived stereotype threat were not statistically different between the control condition ( $M = 2.88, SD = .38$ ) and the combination condition ( $M = 2.98, SD = .27$ ). Additionally, perceived performance pressure was statistically the same in the control ( $M = 3.31, SD = .92$ ), stereotype threat ( $M = 3.70, SD = .94$ ), performance pressure ( $M = 3.32, SD = .94$ ), and combination conditions ( $M = 3.56, SD = 1.08$ ).

### **Hypothesis Testing**

Hypothesis 1 stated that (a) performance would be higher in the control condition than in the stereotype condition, (b) performance would be higher in the control condition than in the performance pressure condition, and (c) performance would be higher in the control condition than in the combination (stereotype threat + performance pressure) condition. A 2x2 Factorial ANOVA revealed that there was a significant main effect for performance pressure [ $F(1, 146) = 5.31, p = .02$ ] such that individuals in the performance pressure conditions ( $M = 6.81, SD = 2.69$ ) had lower scores than individuals who were not in the performance pressure conditions ( $M = 7.90, SD = 2.92$ ). Both the main effect for stereotype threat [ $F(1, 146) = .04, p = .84$ ] and the interaction between stereotype threat and performance pressure [ $F(1, 146) = .96, p = .33$ ] were not statistically significant (See Table 2). Descriptive statistics indicated that the condition with

the highest performance was stereotype threat ( $M = 8.16, SD = 2.89$ ) followed by control ( $M = 7.62, SD = 2.95$ ), performance pressure ( $M = 7.00, SD = 2.60$ ), and combination ( $M = 6.64, SD = 2.79$ ). Relatedly, a multiple regression revealed that both *perceived* stereotype threat and *perceived* performance pressure were significantly related to test performance,  $F(2, 146) = 7.37, p < .01$ . However, the direction of this relationship was opposite to what was expected as both the unstandardized regression coefficients for perceived stereotype threat ( $\beta = .20, p = .02$ ) and perceived performance pressure ( $\beta = .20, p = .02$ ) were positive; this would represent a boost in performance as opposed to a detriment when either stereotype threat or performance pressure were present (See Table 3). Thus, Hypothesis 1b was supported and Hypotheses 1a and 1c were not supported.

Hypothesis 2 stated that (a) competitiveness would moderate the relationship between stereotype threat and performance such that under conditions of stereotype threat, test performance will be higher when individuals report high levels of competitiveness as opposed to low levels of competitiveness and that (b) competitiveness would moderate the relationship between performance pressure and performance such that under conditions of performance pressure, test performance will be higher when individuals report high levels of competitiveness as opposed to low levels of competitiveness. An ANCOVA indicated that the interaction between condition (stereotype threat, performance pressure, combination, or control) and competitiveness was not statistically significant,  $F(3, 141) = .18, p = .91$ . Additionally, a multiple regression was used to investigate all possible two-way and three-way interactions among the following variables: perceived stereotype threat, perceived performance pressure, and competitiveness. The only one of these interaction terms to achieve statistical significance was the interaction between competitiveness and perceived performance pressure ( $\beta = 1.86, p = .01$ ).

However, all interaction terms had tolerance values that were significantly less than .10, which indicated multicollinearity (See Table 4). Thus, Hypotheses 2a and 2b were not supported.

Hypothesis 3 stated that (a) the relationship between stereotype threat and test performance would be partially mediated by test-taking anxiety, (b) the relationship between stereotype threat and test performance would be partially mediated by test-taking motivation, (c) the relationship between performance pressure and test performance would be partially mediated by test-taking anxiety, and (d) the relationship between performance pressure and test performance would be partially mediated by test-taking motivation. *MEDIATE* (Hayes & Preacher, 2011) is an SPSS macro that was used to test for mediation because it can incorporate multiple mediators and multiple independent variables into the same analysis. When using a categorical variable, *MEDIATE* calculates the *relative indirect effects* of each level of the independent variable. Relative indirect effects can be thought of as the effect of being in one group (or set of groups) relative to some reference or comparison group (Hayes & Preacher, 2011). Results indicated that for individuals in the stereotype threat condition, there was a positive indirect effect relative to the control group through anxiety,  $ab = .56$  (95% CI [.10, 1.19]). This indirect effect was in the direction opposite to what was expected, as individuals in the stereotype threat condition experienced lower anxiety than the control group, and that low anxiety led to higher test performance (See Figure 7 and Table 5). *MEDIATE* was also used to test for mediation using perceived stereotype threat and perceived performance pressure as predictors and test-taking anxiety and test-taking motivation as mediators. Results indicated that there was a positive indirect effect of perceived performance pressure on test performance through test-taking motivation,  $ab = .38$  (95% CI [.05, .77]). This indirect effect was also in the direction opposite to what was expected, as individuals who reported feeling a high degree of

performance pressure also reported high levels of motivation, which led to higher test performance (See Table 6). Figure 8 shows the path diagram and coefficients for the indirect effects of perceived stereotype threat and perceived performance pressure through anxiety and motivation. In short, although the analyses produced some results that were statistically significant, the results were not in the hypothesized direction, and thus Hypotheses 3a – 3d were not supported.

Hypothesis 4 stated that (a) the interactive effects of stereotype threat and competitiveness on test performance would be partially mediated by test-taking motivation and (b) the interactive effects of performance pressure and competitiveness on test performance would be partially mediated by test-taking motivation. PROCESS (Hayes, 2012) is an SPSS macro that was used to test for mediated-moderation. It has the ability to test for both moderated-mediation and mediated-moderation using multiple mediators and multiple moderators, as well as dichotomous outcomes. Results indicated that anxiety did not mediate the indirect effect of the interaction of perceived stereotype threat and competitiveness on performance,  $ab = -.34$  (95% CI [-1.26, .53]) or the indirect effect of the interaction of perceived performance pressure and competitiveness on performance,  $ab = .14$  (95% CI [-.27, .58]). Additionally, motivation did not mediate the indirect effect of the interaction of perceived stereotype threat and competitiveness on performance,  $ab = -.20$  (95% CI [-.44, 1.07]) or the indirect effect of the interaction of perceived performance pressure and competitiveness on performance,  $ab = .06$  (95% CI [-.15, .39]) Figures 9 and 10 show the path diagrams and coefficients for these indirect effects. Finally, results indicated that relative to the control condition, competitiveness did not produce a change in performance due to motivation or anxiety for individuals in each of the threat conditions (stereotype threat, performance pressure, combination). Relative indirect effects for stereotype

threat condition \* competitiveness, performance pressure condition \* competitiveness, and combination condition \* competitiveness can be found in Table 7. Figures 11 – 13 show path diagrams and coefficients for each condition (using the control condition as a reference). In short, Hypotheses 4a and 4b were not supported.

## Discussion

### Summary

The primary objectives of the present study were as follows: (a) to compare the effects of stereotype threat and performance pressure directly in order to determine which is the best predictor of performance, (b) to test the moderating effect of competitiveness on the relationship between pressure condition (e.g., stereotype threat or performance pressure) and task performance and (c) to investigate test-related perception variables (e.g., test-taking anxiety, test-taking motivation) as mediators of the interactive effect of pressure condition and competitiveness on task performance. Although there were no specific hypotheses regarding which predictor (e.g., stereotype threat or performance pressure) had the greater impact on performance, it was hypothesized that test performance would be highest in the control condition when both variables were absent and lower in conditions in which either or both of those variables were present. This hypothesis was partially supported as the results indicated a significant main effect for performance pressure; for individuals in either of the two conditions that had performance pressure (e.g., performance pressure condition and combination condition), test performance was significantly lower than individuals who were in either of the remaining two conditions (e.g., stereotype threat condition and control). More interestingly, however, was the finding that *perceptions* of both stereotype threat and performance pressure led to a significant *increase* in test performance. Additionally, these perceptions were statistically similar

across all four conditions. These findings seem to indicate that although the experimental manipulations did not alter perceptions of stereotype threat and performance pressure, those same perceptions can influence test performance in unexpected ways. Furthermore, performance pressure manipulations appear to have a stronger impact on performance than stereotype threat manipulations, although neither seems to affect their respective underlying perception variable.

Hypothesis 2 stated that competitiveness would moderate the relationship between (a) stereotype threat and test performance and (b) performance pressure and test performance. An ANCOVA was used to test this hypothesis so that the analysis could include for the categorical independent variable as well as the continuous moderator variable. Furthermore, when using ANCOVA, moderation can be tested with one interaction term as opposed to three interaction terms, which would have been necessary if multiple regression and dummy-coding were used for this particular analysis. These hypotheses were not supported primarily due to multicollinearity among the interaction terms. This is likely because the interaction terms were essentially tapping the same or similar construct variables. In short, competitiveness did not alter the relationship between stereotype threat and test performance or performance pressure and test performance. Hypothesis 3 stated that test-taking anxiety and test-taking motivation would mediate the relationship between (a) stereotype threat and test performance and (b) performance pressure and performance. Results indicated that for individuals in the stereotype threat condition, there was an indirect effect on test performance through anxiety. Also, there was an indirect effect of perceived performance pressure on test performance through motivation. However, because both of the effects were in the opposite direction (e.g., positive instead of negative), Hypothesis 3 was not supported. Finally, Hypothesis 4 was not supported because anxiety and motivation did not mediate the interactive effects of the predictor variables (e.g., test condition, perceived stereotype

threat, perceived performance pressure) on test performance. There was a significant interaction between race and test condition, but interpretation was somewhat difficult to decipher because there was not an equal representation of all racial groups across test conditions. In general, African American participants performed highest in the control condition, and their performance steadily declined across the stereotype threat, performance pressure, and combination conditions in that order. Hispanic participants performed much better in the control and performance pressure conditions than in the stereotype threat and combination conditions. White participants generally performed the same across all conditions. This finding is surprising because the gender stereotype chosen for this study should have affected all female participants equally, regardless of racial subgroup. Therefore, it was expected that the performance detriments would be roughly the same across all racial subgroups. Figure 14 shows a visual representation of all of the means by test condition across racial subgroups. Additionally, self-reported math ability was not significantly related to perceptions of stereotype threat or test performance in the stereotype threat conditions, meaning that domain identification did not affect the degree to which performance was affected by stereotype threat. Finally, stereotype endorsement was not significantly related to perceptions of stereotype threat or test performance in the stereotype threat conditions, meaning that those individuals who reported that *“I am at a disadvantage on this test due to my gender”* performed no better or worse than those participants who did not endorse the negative stereotype.

### **Practical Implications**

The results of the current study align most closely with those found by Oswald and Harvey (2000-2001), in which stereotype threat activation actually led to an increase in performance. Moderately explicit activation cues were used because they are the cues that are

most likely to be experienced in real-world settings, and they have not previously been shown to result in stereotype reactance (Nguyen & Ryan, 2008); despite this, the results were consistent with a stereotype reactance effect. Test performance was higher in the stereotype threat condition than in the control condition, although that difference did not reach statistical significance. However, perceptions of stereotype threat were *positively* related to test performance. These findings demonstrate that *threatening* situations can be beneficial to the extent that they motivate targeted individuals to perform better than if threat was absent. When the participants in this study were presented with damaging stereotypical information, they strove to perform at a higher level so that they could disconfirm the negative stereotype and maintain their positive group identity.

Additionally, the performance pressure manipulations led to an overall decrease in performance, yet *perceptions* of performance pressure actually led to increases in performance. These perceptions were statistically similar across conditions, meaning that even people in the control and stereotype threat conditions experienced pressure. It appears that the performance pressure *manipulations* are the true explanations for the detriment to performance instead of actual performance pressure. Perceived performance pressure may not have resulted in a decrease in performance because the math ability test did not require enough working memory resources to result in a shortage when dealing with task performance and intrusive thoughts. As a result, the incentives (e.g., money, social evaluation, disconfirming a negative stereotype) served as motivation instead of interference.

The finding that participants in the stereotype threat condition reported lower levels of anxiety and thus the highest levels of performance across all groups was unexpected, and particularly noteworthy. Part of the stereotype threat manipulation included the statement that the



test had produced gender differences. The statement did not include specific information concerning the direction of those differences. For this reason, the female participants may have thought that the gender differences were actually in their favor (e.g., women have performed better than men historically on math ability tests) and thus been given an advantage due to the stereotype instead of a handicap. Additionally, the subtle cues that were used to activate stereotype threat may have had an additive effect upon one another and thus resulted in the same reaction (stereotype reactance) that a blatant cue could have caused.

The discovery that stereotype threat and performance pressure may not always result in a detriment to performance can be considered good news. For some individuals, perceptions of stereotype threat appear to decrease anxiety and perceptions of performance pressure increase motivation. It is important to note that all of the female participants were psychology students, and that the university at which this study took place offers a number of classes and seminars that are designed to address professional issues that women are likely to experience in both academic and career-oriented settings. Because of this, the participants may have been aware of stereotype threat and its effects and therefore resistant to its impact on their performance. Also, it is possible that the stereotype in question may no longer be prevalent. Steele and Aronson (1995) insist that a stereotype must be widely-held in order to elicit threat, and in a world with powerful female role models such as Oprah Winfrey and Sandra Day O'Connor, the idea that women are inferior to men with respect to academics may have fallen out of favor. Recent research has demonstrated that although men report more positive attitudes towards math, actual gender differences in math ability are both small and negligible (Else-Quest, Hyde, & Linn, 2010).

### **Limitations and Future Research**

To ensure that stereotype threat was truly activated, it may have been useful to include a comparison sample of male participants to see if there were differences across gender. There remains the possibility that any participants, regardless of gender, could have achieved similar performance across the stereotype threat and performance conditions. Additionally, it may have been useful to test female participants in the presence of males other who were taking the same test. Previous research (e.g. Inzlicht & Ben-Zeev, 2000; 2003; Sekaquaptewa & Thompson, 2002; 2003) supports the proposition that stereotype salience and subsequent threat increases in the presence of majority subgroup members. Finally, while the math ability test that was chosen produced a substantial amount of variability among participants, a more difficult test may have intensified the detrimental effects of both stereotype threat and performance pressure.

Future research concerning stereotype threat and performance pressure in women should include a more difficult measure of math ability as well as a sample of male participants to determine if these same affects apply when investigating these variables with the majority subgroup instead of the minority subgroup. Also, using participants from a non-academic setting may provide results that are more consistent with previous research because those participants may be less likely to have access to the classes and services that many universities offer that are designed to prepare women threatening professional situations.

Lastly, future research should investigate stereotype reactance effects, particularly with respect to the gender stereotype chosen for this study. The results were consistent with a stereotype reactance effect with participants in the threat condition performing the best of all groups. Blatant cues that have been demonstrated to cause stereotype reactance are typically less volatile than subtle cues and, in the case of this study, may have resulted in a larger effect (e.g., higher test performance) in the stereotype threat condition.

## **Conclusion**

Based on the results of this study, it appears that although stereotype threat and performance pressure have historically resulted in similar performance-based outcomes, performance pressure manipulations (e.g., monetary incentives, peer pressure) appear to have more of an impact with respect to math test performance. Also, anxiety and motivation appear to be key factors in the effects of stereotype threat and performance pressure respectively on test performance, although in this case those effects improved performance instead of hindered it. The results also give us hope that widely-held negative stereotypes may be dissipating or gone altogether. Finally, the results indicate that threatening and stressful may not always be bad. There is hope for individuals who find themselves in stressful situations that involve salient consequences for performance such as confirming a negative stereotype or missing out on a cash reward, because perceptions of threat and pressure can result in a performance advantage.

## References

- Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science, 5*, 385-390.
- Aronson, J., Lusting, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When White men can't do math: Necessary and insufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35*, 29-46.
- Arthur, W. Jr., & Doverspike, D. (2005). Achieving diversity and reducing discrimination in the workplace through human resource management practices: Implications of research and theory for staffing, training, and rewarding performance. In R. L. Dipboye, & A. Colella (Eds), *Discrimination at work: The psychological and organizational bases*. (pp. 305-327). Mahwah, NJ: LEA.
- Arvey, R. D., Strickland, W., Drauden, G., & Martin, C. (1990). Motivational components of test taking. *Personnel Psychology, 43*, 695-716.
- Bailey, A. A. (2004). *Effects of stereotype threat on females in math and science fields: An investigation of possible mediators and moderators of the threat-performance relationship*. Unpublished doctoral dissertation, Georgia Institute of Technology.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Bargh, J. A. (1997). The automaticity of everyday life. In R. S. Wyer Jr. (Ed.) *Advances in social cognition* (Vol. 10, pp. 1-61). Mahwah, NJ: Erlbaum.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.
- Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology, 46*, 610-620.
- Beilock, S. L. (2008). Math performance in stressful situations. *Current Directions in Psychological Science, 17*, 339-343.
- Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General, 130*, 701-725.
- Berry, III, R. Q. (2003). Mathematics standards, cultural styles, and learning preferences. The plight and the promise of African American students. *The Clearing House, 76*, 244-249.

- Blanton, H., Christie, C., & Dye, M. (2002). Social identity versus reference frame comparisons: The moderating role of stereotype endorsement. *Journal of Experimental Social Psychology, 38*, 253-267.
- Blanz, M., Mummendey, A., Mielke, R., & Klink, A. (1998). Strategic responses to negative social identity: An empirical systematization of field data. *European Journal of Social Psychology, 28*, 697-729.
- Bobko, P., Roth, P. L., & Potosky, D. (1999). Derivation and implications of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. *Personnel Psychology, 52*, 561-589.
- Bornholt, J. J., Goodnow, J. J., & Cooney, G. H. (1994). Influences of gender stereotypes on adolescents' perceptions of their own achievement. *American Educational Research Journal, 31*, 675-692.
- Brown, R. P., & Day, E. A. (2006). The difference isn't Black and White: Stereotype threat and the race gap on Raven's Advanced Progressive Matrices. *Journal of Applied Psychology, 91*, 979-985.
- Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L., & Latinotti, S. (2003). Stereotype threat: The effect of expectancy on performance. *European Journal of Social Psychology, 33*, 267-285.
- Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology, 27*, 703-722.
- Ceci, S.J., & Williams, W. M. (1997). Schooling, intelligence, and income. *American Psychologist, 52*, 1051-1058.
- Chan, D., & Schmitt, N. (1997). Video-based versus paper-and-pencil method of assessment in situational judgment tests: Subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology, 82*, 143-159.
- Crocker, J., Major, B., & Steele, C. M. (1998). Social stigma. In D.T. Gilbert, S.T. Fiske, & G. Lindzey (Eds.), *The Handbook of Social Psychology* (4<sup>th</sup> edition, Vol. II, pp.504-553). Boston: The McGraw-Hill Companies.
- Crocker, J., Luhtanen, R., Blaine, B., & Broadnax, S. (1994). Collective self-esteem and psychological well-being among White, Black, and Asian college students. *Personality and Social Psychology Bulletin, 20*, 506-513.
- Cullen, M. J., Hardison, C. M., & Sackett, P. R. (2004). Using SAT-grade and ability-job performance relationships to test predictions derived from stereotype threat theory. *Journal of Applied Psychology, 89*, 220-230.

- Davis, C., Aronson, J., Salinas, M. (2006). Shades of threat: Racial identity as a moderator of stereotype threat. *Journal of Black Psychology*, 32, 399-417.
- DeCaro, M. S., Rotar, K. E., Kendra, M. S., & Beilock, S. L. (2010). Diagnosing and alleviating the impact of performance pressure on mathematical problem solving. *Quarterly Journal of Experimental Psychology*, 63, 1619–1630.
- Doojse, B., & Ellemers, N. (1997). Stereotyping under threat: The role of group identification. In R. Spears, P. J. Oakes, N. Ellemers, & S. A. Haslam (Eds.), *The social psychology of stereotyping and group life* (pp. 257-272). Cambridge, MA: Blackwell.
- Doverspike, D., Taylor, M. A., & Arthur, W. E., Jr. (1999). *Affirmative action: A psychological perspective*. New York: Nova Science Publishers, Inc.
- Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly*, 11, 135-172.
- Eccles, J. S., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self and task perceptions during elementary school. *Child Development*, 64, 830-847.
- Edwards, B. D. (2004). *An examination of factors contributing to a reduction in race-based subgroup differences on a constructed response paper-and-pencil test of achievement*. Unpublished doctoral dissertation, Texas A&M University.
- Edwards, B. D., & Arthur, W. A. (2007). An examination of factors contributing to a reduction in subgroup differences on a constructed-response paper-and-pencil test of scholastic achievement. *Journal of Applied Psychology*, 92, 794-801.
- Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics achievement, attitudes, & affect: A meta-analysis. *Psychological Bulletin*, 136, 103-127.
- Fennema, E., & Peterson, P. (1985). "Autonomous learning behavior: A possible explanation of gender-related differences in mathematics." In L. C. Wilkinson & C. B. Marrett (Eds.), *Gender-related Differences in Classroom Interactions*. Academic Press, pp. 17-35.
- Ford, T. E., Ferguson, M. A., Brooks, J. L., & Hagadone, K. M. (2004) Coping sense of humor reduces effects of stereotype threat on women's math performance. *Personality and Social Psychology Bulletin*, 30, 643-653.
- Gottfredson, L. S. (1988). Reconsidering fairness: A matter of ethical and social priorities. *Journal of Vocational Behavior*, 33(3), 293-319.
- Gough, H. G. (1987). *Manual for the California Psychological Inventory*. Palo Alto, CA: Consulting Psychology Press.

- Griffin-Pierson, S. (1990). The Competitiveness Questionnaire: A measure of two components of competitiveness. *Measurement and Evaluation in Counseling and Development*, 23, 108-115.
- Guajardo, G. A. (2005). *Modifying stereotype relevance and altering affect attributions to reduce performance suppression on cognitive ability selection tests*. Unpublished master's thesis, Northern Illinois University.
- Halpern, D. F. (1986). Sex differences in cognitive abilities. Hillsdale, NJ: Erlbaum.
- Hardy, L., Mullen, R., & Jones, G. (1996). Knowledge and conscious control of motor actions under stress. *British Journal of Psychology*, 87, 621–636.
- Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White paper]. Retrieved from <http://www.afhayes.com/public/process2012.pdf>
- Hayes, A. F., & Preacher, K. J. (2011). *Indirect and direct effects of a multicategorical causal agent in statistical mediation analysis*. Manuscript submitted for publication.
- Helmreich, R. L., & Spence, J. T. (1978). Work and Family Orientation Questionnaire: An objective instrument to assess components of achievement motivation and attitudes toward family and career. *JSAS Catalog of Selected Documents in Psychology*, 8, 35.
- Helms, J. E. (1990). *Black and White racial identity: Theory, research, and practice*. New York: Greenwood.
- Hernstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.
- Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science*, 11, 365-371.
- Inzlicht, M., & Ben-Zeev, T. (2003). Do high-achieving female students underperform in private? The implications of threatening environments on intellectual processing. *Journal of Educational Psychology*, 95, 796-805.
- Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, 73, 509-527.
- Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, 5, 602-619.

- Kray, L. J., Thompson, L., & Galinsky, A. D. (2001). Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. *Journal of Personality and Social Psychology, 80*, 942-958.
- Kuiper, N. A., McKenzie, S. D., & Balanger, K. A. (1995). Cognitive appraisals and individual differences in sense of humor: Motivational and affective implications. *Personality and Individual Differences, 19*, 359-372.
- Kuncel, N. R., Hezlett, S. A., & Ones, D. (2001). A comprehensive meta-analysis of the predictive validity of the graduate records examinations: Implication for graduate student selection and performance. *Psychological Bulletin, 127*, 162-181.
- Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and Social Psychology Bulletin, 18*, 302-318.
- Marsh, H. W., Trautwein, U., Ludtke, O., Koller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*, 397-416.
- Marsh, H. W., & Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. *Journal of Educational Psychology, 89*, 41-54.
- Martin, R. A. (1996). The Situational Humor Response Questionnaire (SHRQ) and Coping Sense of Humor Scale (CHS): A decade of research findings. *Humor: International Journal of Humor Research, 9*, 251-272.
- Mayer, D. M., & Hanges, P. J. (2003). Understanding the stereotype threat effect with "culture free" tests: An examination of its mediators and measurement. *Human Performance, 16*, 207-230.
- McKay, P., Doverspike, D., Bowen-Hilton, D., & Martin, Q. (2002). Stereotype effects on Raven's Advanced Progressive Matrices scores of African-Americans. *Journal of Applied Social Psychology, 34*, 767-787.
- McKay, P. F. (1999). *Stereotype threat and its effect on the cognitive ability test performance of African-Americans: The development of a theoretical model*. Unpublished doctoral dissertation, University of Akron, Akron, Ohio.
- Mummendey, A., Kessler, T., Klink, A., & Mielke, R. (1999). Strategies to cope with negative social identity: Predictions by social identity and relative deprivation theory. *Journal of Personality and Social Psychology, 76*, 229-245.
- Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*, 77-101.



- Norem, J. K., & Cantor, N. (1986). Defensive pessimism: “Harnessing” anxiety as motivation. *Journal of Personality and Social Psychology*, *51*, 1208-1217.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math  $\neq$  me. *Journal of Personality and Social Psychology*, *83*(1), 44-59.
- Nguyen, H. D., O’Neal, A., & Ryan, A. M. (2003). Relating test-taking attitudes and skills and stereotype threat effects to the racial gap in cognitive ability test performance. *Human Performance*, *16*, 261-293.
- Nguyen, H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology*, *93*, 1314-1334.
- Oswald, D. L., & Harvey, R. D. (2000-2001). Hostile environments, stereotype threat, and math performance among undergraduate women. *Current Psychology: Developmental, Learning, Personality, Social*, *19*, 338-356.
- Pellegrini, A. V. (2005). *The impact of stereotype threat on intelligence testing in Hispanic females*. Unpublished doctoral dissertation, Carlos Albizu University.
- Perry, S. P., & Skitka, L. J. (2009). Making lemonade? Defensive coping style moderates the effect of stereotype threat on women’s math test performance. *Journal of Research in Personality*, *43*, 918-920.
- Ployhart, R. E., Ziegert, J. C., & McFarland, L. A. (2003). Understanding racial differences on cognitive ability tests in selection contexts: An integration of stereotype threat and applicant reactions research. *Human Performance*, *16*, 231-259.
- Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven’s Progressive Matrices and Vocabulary Scales*. Oxford, England: Oxford Psychologists Press.
- Reynolds, C., & Brown, R. (1984). *Perspectives on bias in mental testing*. New York: Plenum.
- Roth P. L., Bevier, C. A., Bobko, P., Switzer, F. S., III, & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology*, *54*, 297-330.
- Rudman, L. A., & Fairchild, K. (2004). Reactions to counterstereotypic behavior: The role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology*, *87*, 157–176.
- Rushton, J. P., & Ankey, C. D. (1996). Brain size and cognitive ability: Correlations with age, sex, social class, and race. *Psychonomic Bulletin and Review*, *3*, 21-36.

- Ryan, A. M., & Greguras, G. J. (1998). Life is not multiple choice: Reactions to the alternatives. In Hakel, M. D. (Ed.). *Beyond multiple-choice: Evaluating alternatives to traditional testing for selection*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Saccuzzo, D. P., & Johnson, N. E. (1995). Traditional psychometric tests and proportionate representation: An intervention and program evaluation study. *Psychological Assessment*, 7, 183-194.
- Sackett, P. R., Hardison, C. M., & Cullen, M. J. (2004). On interpreting stereotype threat as accounting for African American-White differences on cognitive tests. *American Psychologist*, 59, 7-13.
- Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). High-stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative-action world. *American Psychologist*, 56, 302-318.
- Sawyer, T. P., & Hollis-Sawyer, L. A. (2003, April). Stereotype threat and face validity effects on cognitive test performance. Poster session presented at the annual meeting of the Society for Industrial Organizational Psychology (SIOP), Orlando, FL.
- Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology*, 38, 194-201.
- Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, 85, 440-452.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-273.
- Schmitt, N., Clause, C. S., & Pulakos, E. D. (1996). Subgroup differences associated with different measures of some common job relevant constructs. *International Review of Industrial and Organizational Psychology*, 11, 115-139.
- Sekaquaptewa, D., & Thompson, M. (2002). The differential effects of solo status on members of high- and low-status groups. *Personality and Social Psychology Bulletin*, 28, 694-707.
- Sekaquaptewa, D., & Thompson, M. (2003). Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal of Experimental Social Psychology*, 39, 68-74.
- Smither, J. W., Reilly, R. R., Millsap, R. E., Pearlman, K., & Stoffey, R. W. (1993). Applicant reactions to selection procedures. *Personnel Psychology*, 46, 49-76.

- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological Methodology 1982* (pp. 290-312). Washington DC: American Sociological Association.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math test performance. *Journal of Experimental Social Psychology, 35*, 4-28.
- Spencer, S. J., Fein, S., Strahan, E. J., & Zanna, M. P. (2005). The role of motivation in the unconscious: How our motives control the activation of our thoughts and shape our actions. In K. D. Williams & J. P. Forgas (Eds.), *Social motivation: Conscious and unconscious processes* (pp. 113-129). New York: Cambridge University Press.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African-Americans. *Journal of Personality and Social Psychology, 69*, 797-811.
- Steele, C. M., & Aronson, J. (1998). Stereotype threat and the test performance of academically successful African Americans. In Jencks, C., & Meredith, P. *The Black-White Test Score Gap* (pp. 401-427). Washington D.C.: Brookings Institution.
- Steele, C. M., & Aronson, J. (2004). Stereotype threat does not live by Steele and Aronson (1995) alone. *American Psychologist, 59*, 47-48.
- Steele, C. M., & Davies, P. G. (2003). Stereotype threat and employment testing: A commentary. *Human Performance, 16*, 311-326.
- Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype threat and social identity threat. *Advances in Experimental Social Psychology, 34*, 379-440.
- Stetsenko, A., Little, T. D., Gordeeva, T., Grasshof, M., & Oettingen, G. (2000). Gender effects in children's beliefs about school performance: A cross-cultural study. *Child Development, 71*, 517-527.
- Stricker, L. J. (1998). *Inquiring about examinees' ethnicity and sex: effects on AP Calculus AB Examination performance* (College Board Rep. No. 98-1; ETS Research Rep. No. 98-5). New York: college Entrance Examination Board.
- Stricker, L. J., & Ward, W. C. (1998). *Inquiring about examinees; ethnicity and sex: Effects on computerized placement tests performance*. (College Board Rep. 98-2; ETS Research Rep. No. 98-9). New York: College Entrance Examination Board.
- Stryker, S., & Serpe, R.T. (1982). Commitment, identity salience and role behavior. In W. Ickes & E. Knowles (Eds.), *Personality, roles, and social behavior* (pp. 199-218). New York: Springer-Verlag.

- Stryker, S., & Serpe, R.T. (1994). Identity salience and psychological centrality: Equivalent, overlapping, or complementary concepts? *Social Psychology Quarterly*, 57, 16-36.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-47). Monterey, CA: Brooks/Cole.
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (2nd ed., (pp. 7-24). Chicago: Nelson-Hall.
- Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbretton, A., Freedman-Doan, C., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology*, 89, 451-469.
- Yovetick, N. A., Dale, J. A., & Hudak, M. A. (1990). Benefits of humor in reduction of threat-induced anxiety. *Psychological Reports*, 66, 51-58.
- Ziegert, J. C., Ployhart, R. E., & McFarland, L. A. (2002). *Perceived stereotype threat: Development of a self-report scale*. Paper presented at the annual conference of the Society of Industrial and Organizational Psychology, Toronto, Canada.

## APPENDIX A

Table 1

*Means, Standard Deviations, Intercorrelations, for All Test Variables.*

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6
1. Math Test Score	7.41	2.86						
2. Perceived ST	2.94	0.37	.23**					
3. Perceived PP	3.48	0.98	.25**	.18*				
4. Motivation	3.86	0.56	.31**	.23**	.61**			
5. Anxiety	2.87	0.70	-.38**	.01	-.03	-.03		
6. Competitiveness	3.41	0.44	.04	.22**	.23**	.29**	.16*	

Note: \* $p < .05$ ; \*\* $p < .01$ . Perceived ST = perceived stereotype threat; Perceived PP = perceived performance pressure.

Table 2

*2x2 Factorial ANOVA: Stereotype Threat and Performance Pressure.*

Source	<i>df</i>	<i>F</i>	$\eta^2$	<i>p</i>
(A) Stereotype Threat	1	.04	.00	.84
(B) Performance Pressure	1	5.31	.04	.02
A*B (interaction)	1	.96	.01	.33
Error	146			

Table 3

*Multiple Regression: Perceived Stereotype Threat and Perceived Performance Pressure.*

Variable	<i>B</i>	<i>SE B</i>	$\beta$
(Constant)	.97	1.86	
Perceived Stereotype Threat	1.51	.62	.20*
Perceived Performance Pressure	.59	.24	.20*

Note:  $R^2 = .09$ . \* $p < .05$



Table 4

*Hypothesis 2 Excluded Variables.*

Excluded Variable	$\beta$	$p$	Tolerance
ST * Competitiveness	.46	.62	.01
PP * Competitiveness	1.86	.01	.01
ST * PP	.22	.79	.01
ST * PP * Competitiveness	.79	.11	.03

Note: ST = perceived stereotype threat; PP = perceived performance pressure.

Table 5

*Indirect Effects of Condition on Test Performance through Anxiety and Motivation.*

Mediator	Indirect Effect	Standard Error	Lower Limit 95% C.I.	Upper Limit 95% C.I.
Stereotype Threat Condition				
Anxiety	.56	.28	.10	1.19
Motivation	.25	.23	-.17	.76
Performance Pressure Condition				
Anxiety	.27	.26	-.24	.81
Motivation	.15	.19	-.20	.58
Combination Condition				
Anxiety	.23	.27	-.36	.75
Motivation	.35	.22	-.04	.81

Table 6

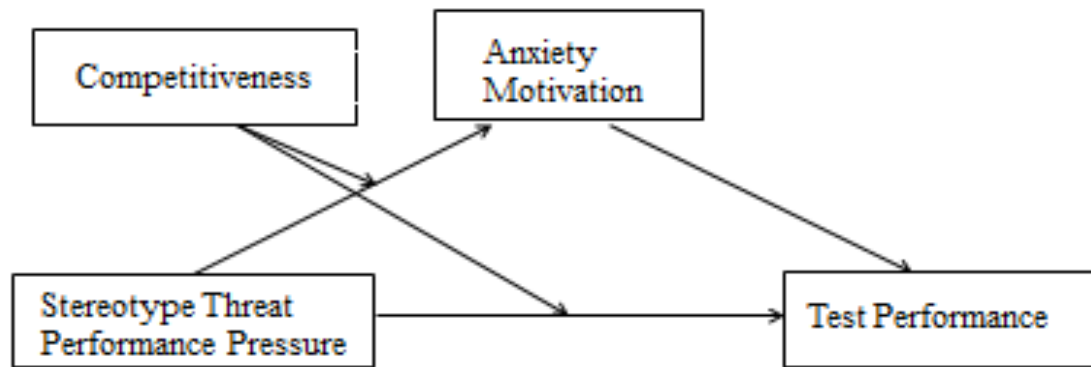
*Indirect Effects of Perceived Stereotype Threat and Perceived Performance Pressure through Anxiety and Motivation.*

Mediator	Indirect Effect	Standard Error	Lower Limit 95% C.I.	Upper Limit 95% C.I.
Perceived Stereotype Threat				
Anxiety	-.01	.24	-.46	.49
Motivation	.22	.18	-.04	.66
Perceived Performance Pressure				
Anxiety	.02	.11	-.20	.21
Motivation	.38	.19	.05	.77

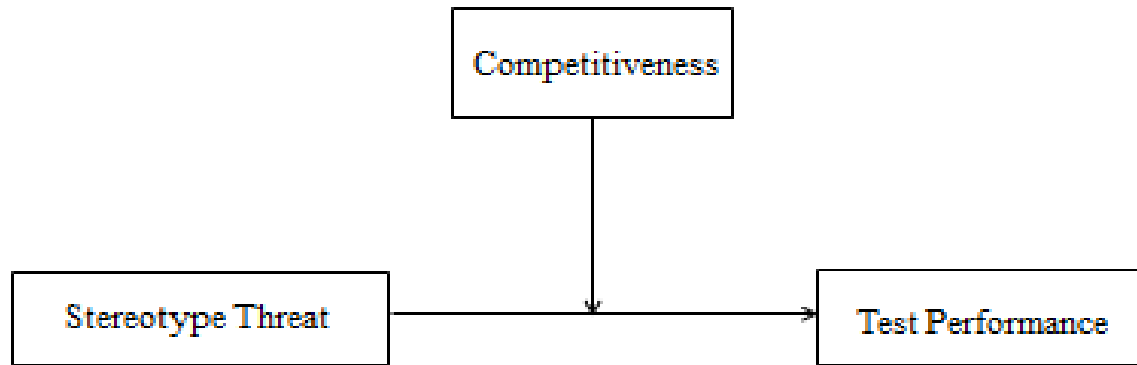
Table 7

*Indirect Effects of Stereotype Threat Condition \* Competitiveness, Performance Pressure Condition \* Competitiveness, and Combination Condition \* Competitiveness on Performance through Anxiety and Motivation.*

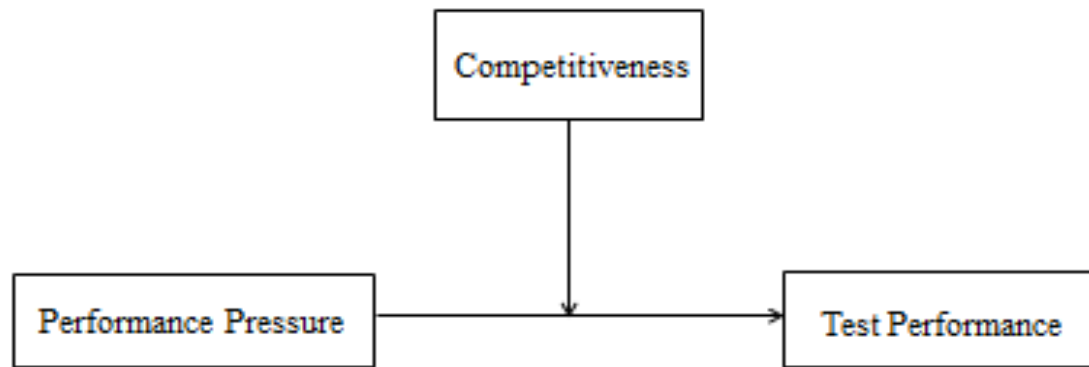
Mediator	Indirect Effect	Standard Error	Lower Limit 95% C.I.	Upper Limit 95% C.I.
Stereotype Threat Condition * Competitiveness				
Anxiety	-.03	.45	-.96	.82
Motivation	.26	.46	-.64	1.14
Performance Pressure Condition * Competitiveness				
Anxiety	.60	.42	-.15	1.55
Motivation	.04	.41	-.77	.89
Combination Condition * Competitiveness				
Anxiety	.25	.58	-.81	1.55
Motivation	-.19	.52	-1.25	.79



*Figure 1.* Conceptual Model of Stereotype Threat and Performance Pressure Effects (Adapted from Hayes, 2012).



*Figure 2.* Conceptual Model of Moderation Effect of Competitiveness on the Relationship between Stereotype Threat and Test Performance



*Figure 3.* Conceptual Model: Moderation Effect of Competitiveness on the Relationship between Performance Pressure and Test Performance

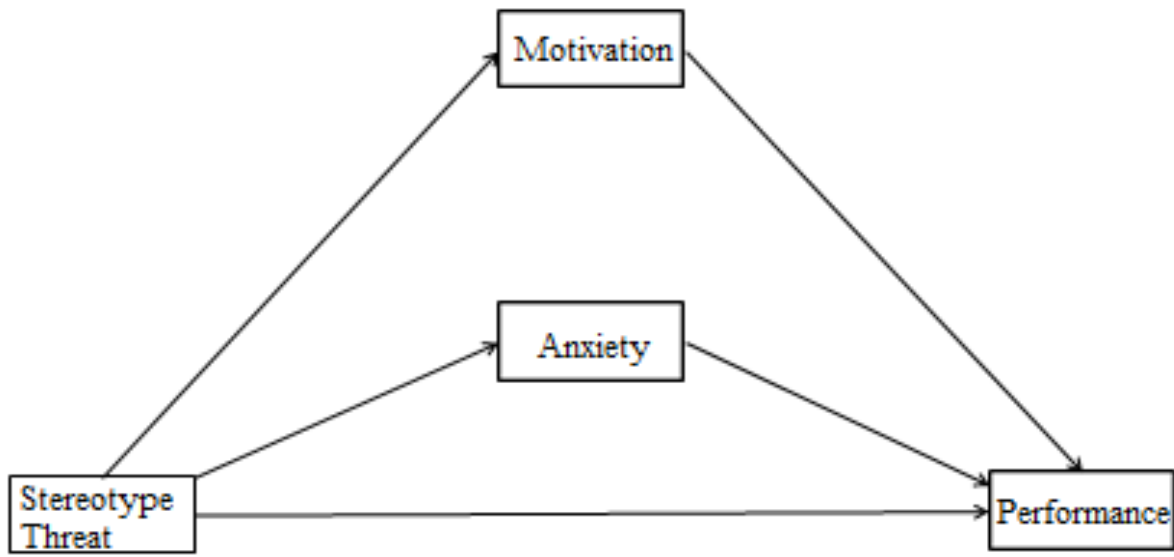
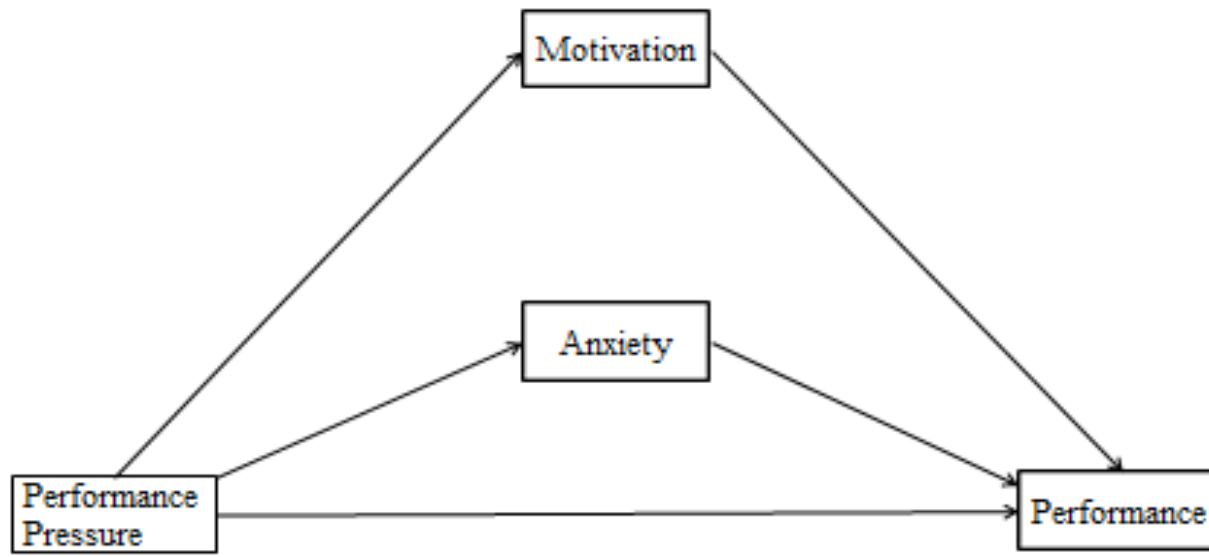


Figure 4. Conceptual Model: Indirect Effect of Stereotype Threat on Test Performance through Anxiety and Motivation.





*Figure 5.* Conceptual Model: Indirect Effect of Performance Pressure on Test Performance through Anxiety and Motivation.

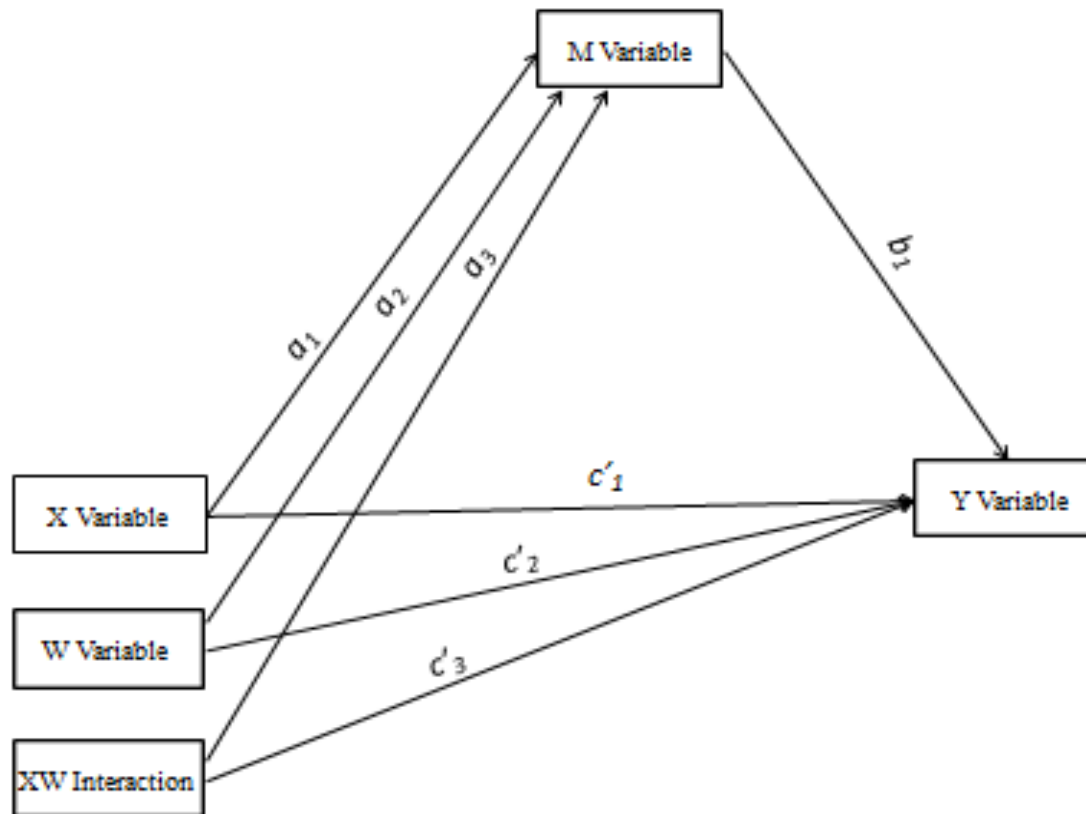


Figure 6. Statistical Model of Mediated Moderation.

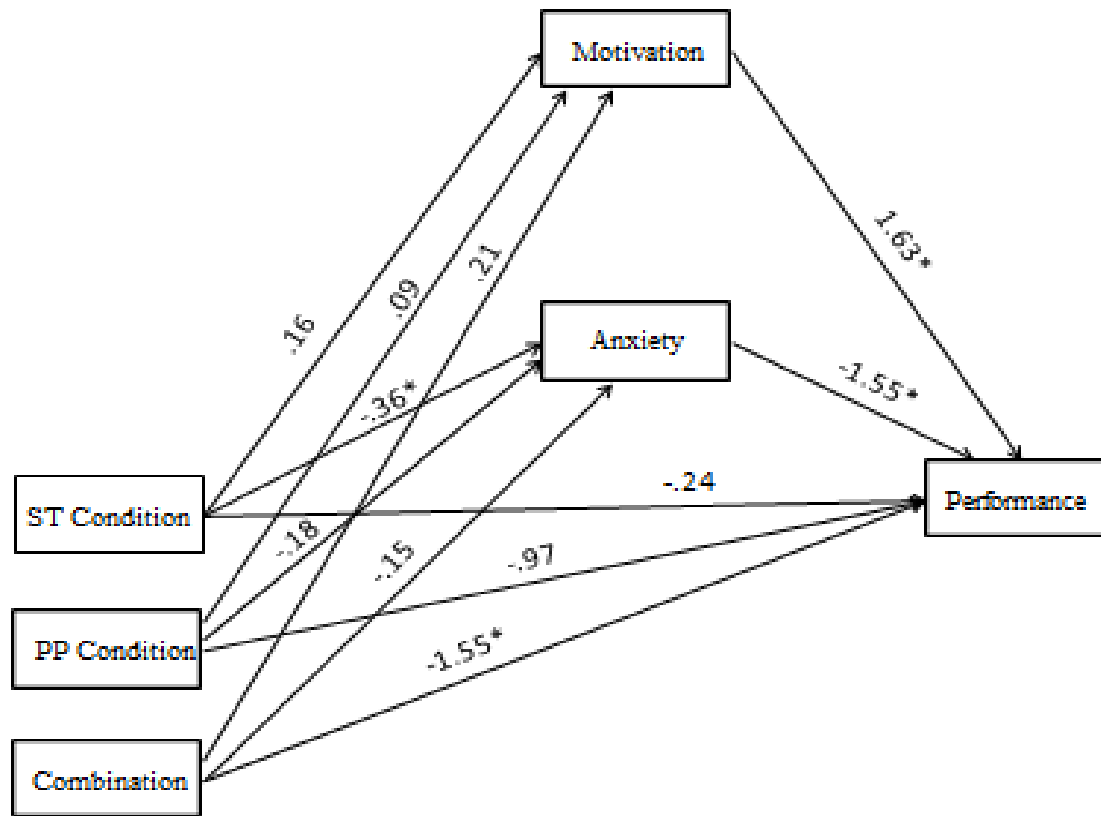


Figure 7. Indirect Effect of Condition through Motivation and Anxiety.

Note: \* $p < .05$ ; \*\* $p < .01$ . ST Condition = Stereotype Threat Condition; PP Condition = Performance Pressure Condition. All path coefficients are unstandardized.

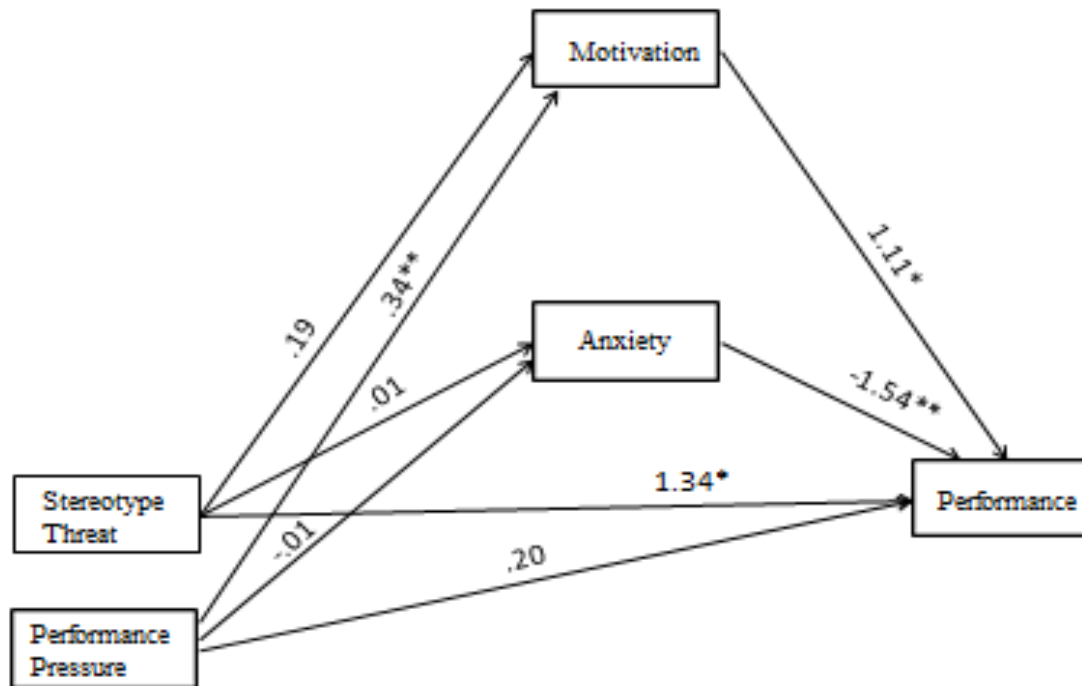


Figure 8. Indirect Effects of Stereotype Threat and Performance Pressure on Performance through Anxiety and Motivation

Note: \* $p < .05$ ; \*\* $p < .01$ . Stereotype Threat = Perceived Stereotype Threat; Performance Pressure = Perceived Performance Pressure. All path coefficients are unstandardized.

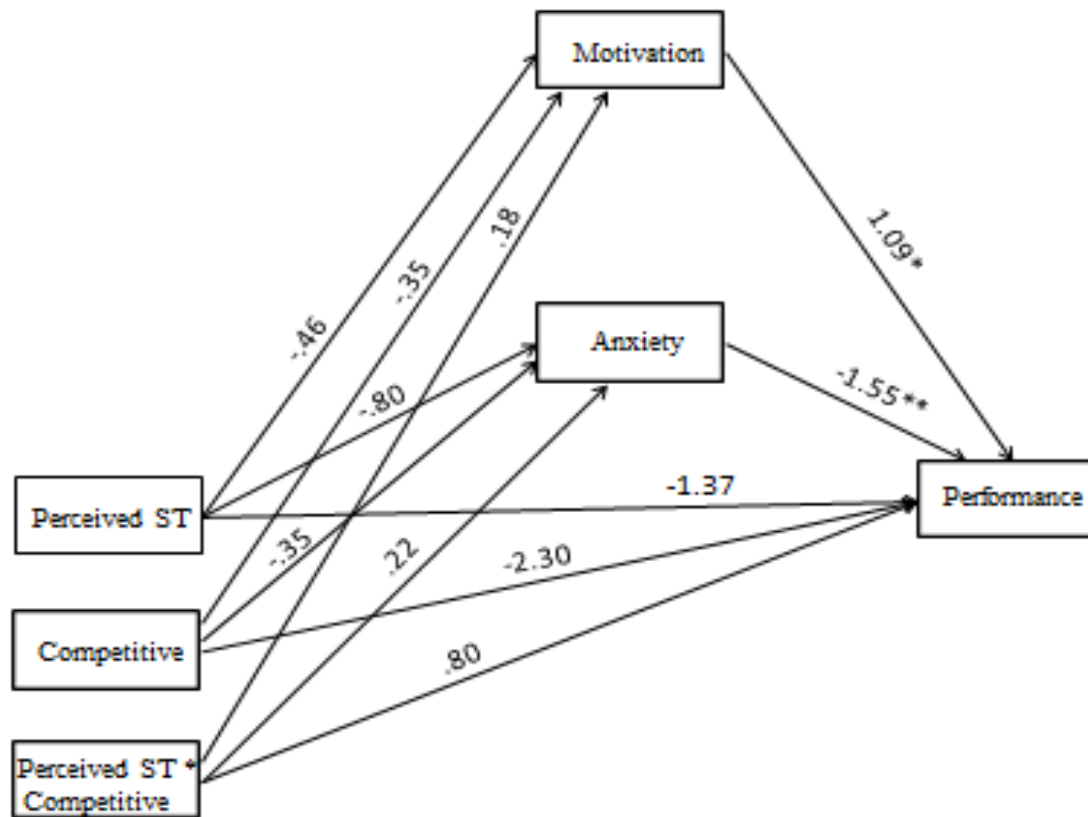


Figure 9. Indirect Effect of the Interaction of Perceived Stereotype Threat and Competitiveness on Performance through Anxiety and Motivation

Note: \* $p < .05$ ; \*\* $p < .01$ . Perceived ST = Perceived Stereotype Threat; Competitive = Competitiveness. All path coefficients are unstandardized.

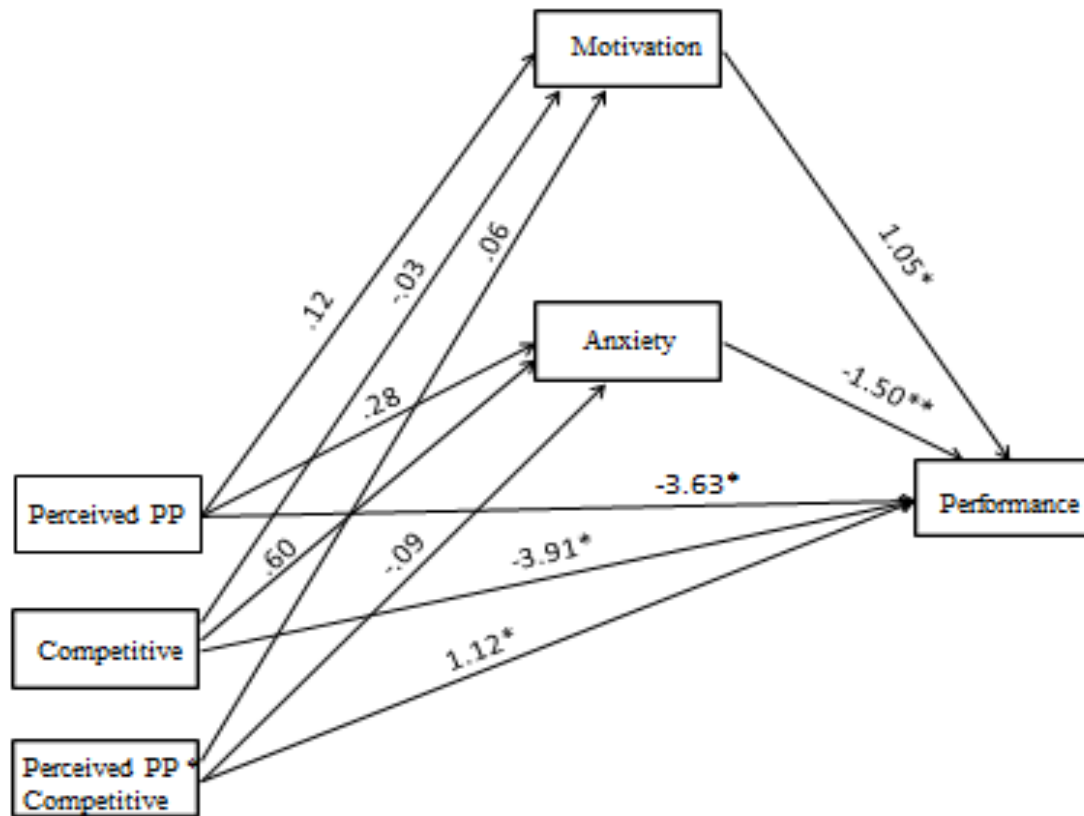


Figure 10. Indirect Effect of the Interaction of Perceived Performance Pressure and Competitiveness on Performance through Anxiety and Motivation.

Note: \* $p < .05$ ; \*\* $p < .01$ . Perceived PP = Perceived Performance Pressure; Competitive = Competitiveness. All path coefficients are unstandardized.

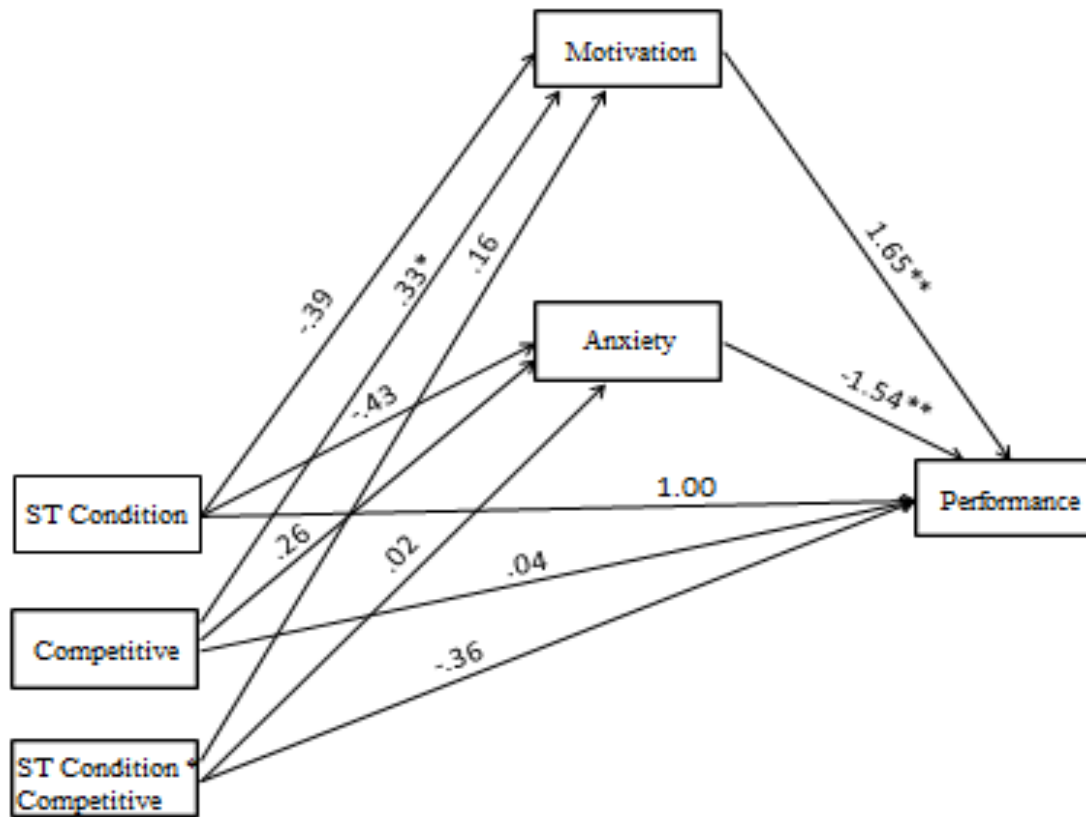


Figure 11. Indirect Effect of the Interaction of Stereotype Threat Condition and Competitiveness on Performance through Anxiety and Motivation.

Note: \* $p < .05$ ; \*\* $p < .01$ . ST Condition = Stereotype Threat Condition; Competitive = Competitiveness. All path coefficients are unstandardized.

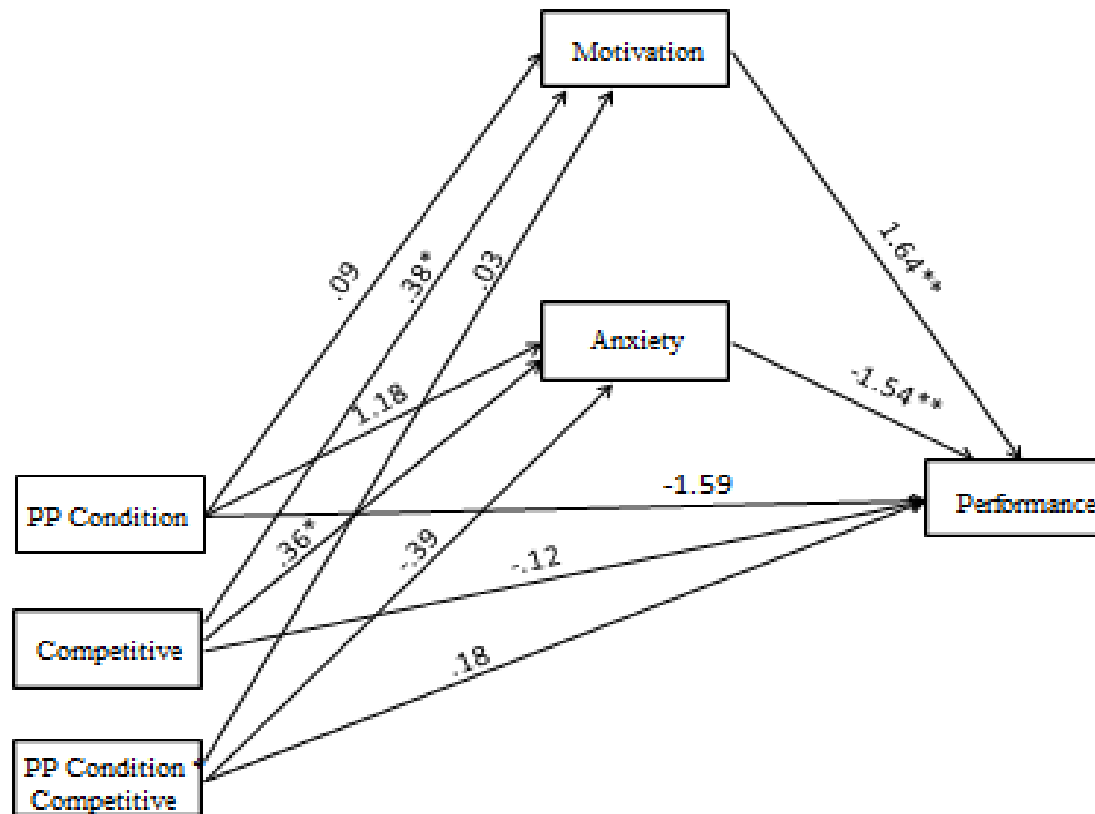


Figure 12. Indirect Effect of the Interaction of Performance Pressure Condition and Competitiveness on Performance through Anxiety and Motivation.

Note: \* $p < .05$ ; \*\* $p < .01$ . PP Condition = Performance Pressure Condition; Competitive = Competitiveness. All path coefficients are unstandardized.



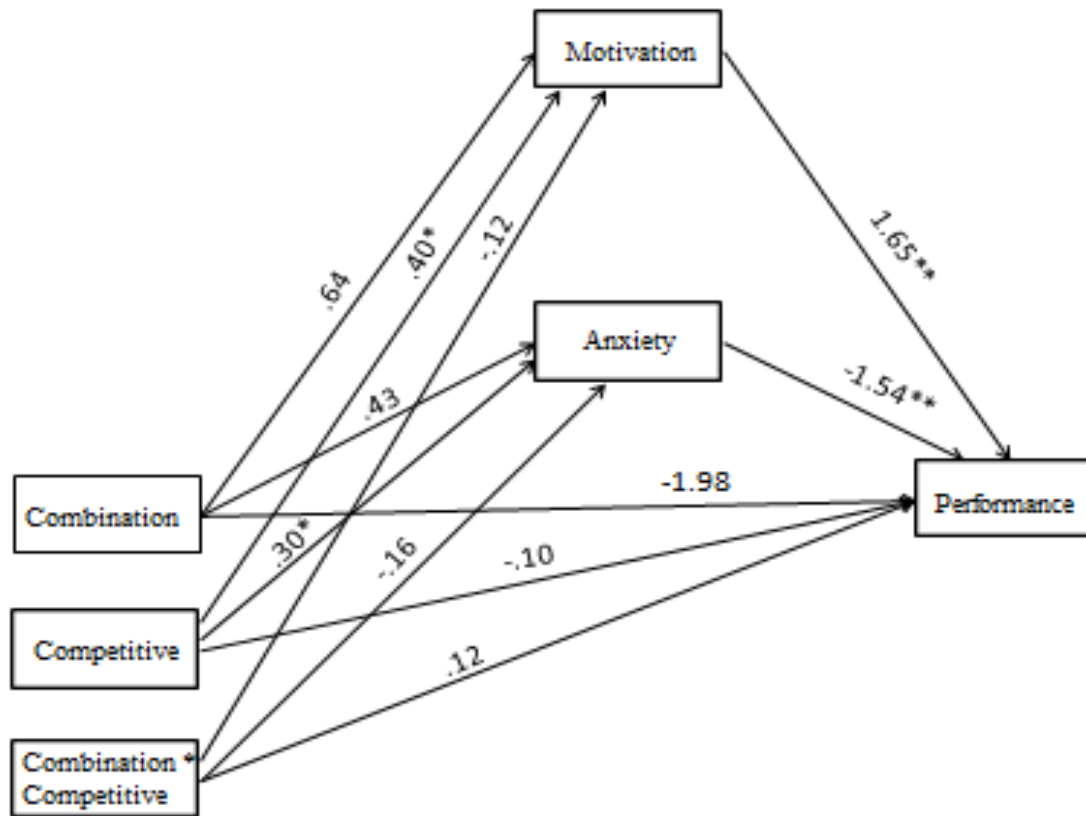


Figure 13. Indirect Effect of the Interaction of Combination Condition and Competitiveness on Performance through Anxiety and Motivation.

Note: \* $p < .05$ ; \*\* $p < .01$ . Competitive = Competitiveness. All path coefficients are unstandardized.

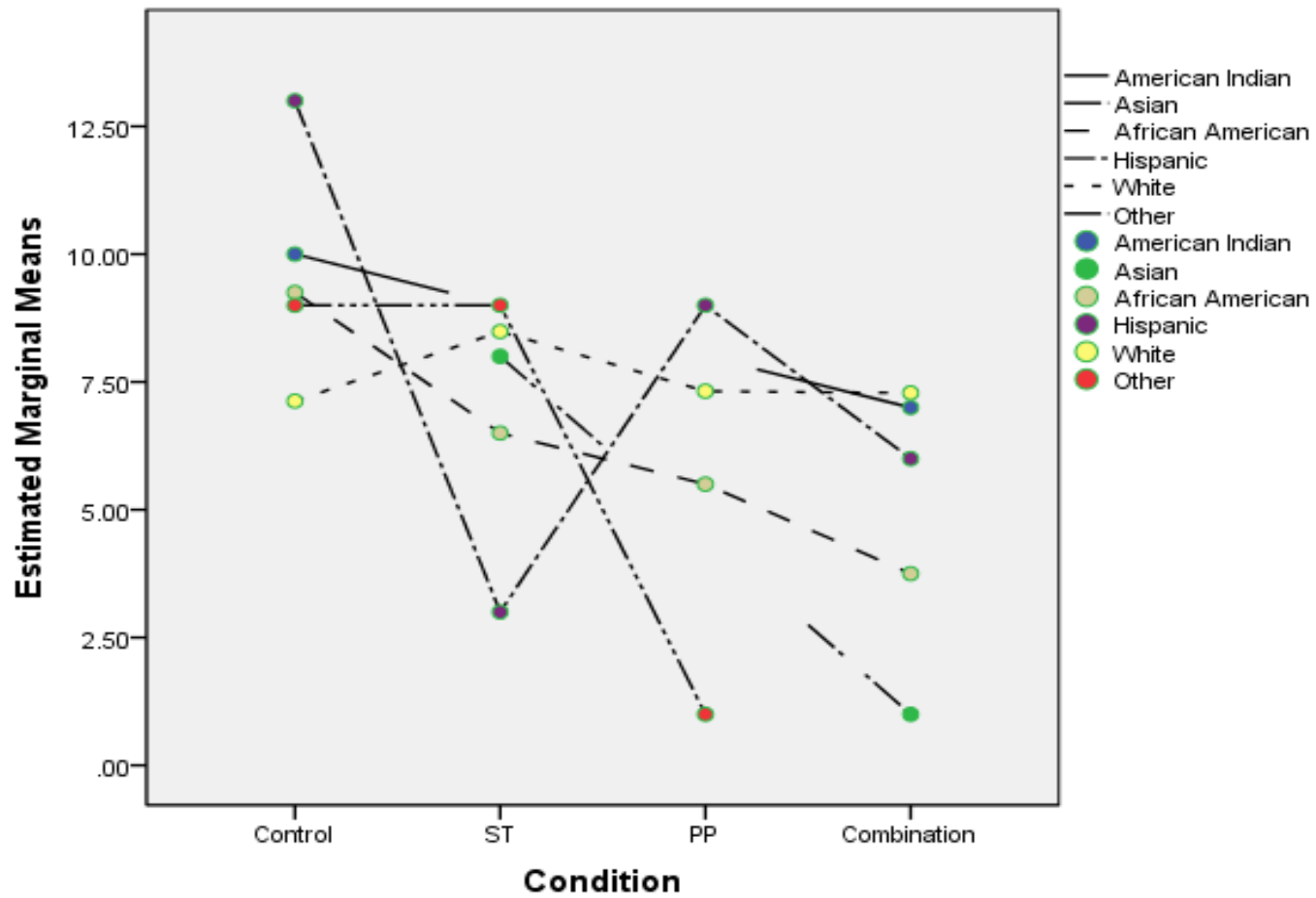


Figure 14. Average Performance by Racial Subgroup Across Conditions.

## APPENDIX B

**PERSONAL ATTITUDES AND BELIEFS**

**DIRECTIONS**

The following items ask you to rate your general attitudes, beliefs, and opinions about gender issues. Read each statement carefully and give your honest feelings about the beliefs and attitudes expressed. Indicate the extent to which you agree with each item by circling the number on the accompanying scale that corresponds to your answer.

<b>Strongly Disagree</b>	<b>Disagree</b>	<b>Neither Disagree nor Agree</b>	<b>Agree</b>	<b>Strongly Agree</b>
1	2	3	4	5

<i><b>Perceived Stereotype Threat</b></i>				
1. People of my gender do significantly better on intelligence tests.	1	2	3	4 5
2. I think others believe that my gender determines how well I do on intelligence tests.	1	2	3	4 5
3. I actually have an advantage on intelligence tests due to my gender.	1	2	3	4 5
4. I am at a disadvantage on intelligence tests due to my gender.	1	2	3	4 5
5. The test may have been easier for people of my gender.	1	2	3	4 5
6. The experimenter expected me to do poorly on the test because of my gender.	1	2	3	4 5
7. In college classes, people of my gender often face biased evaluations.	1	2	3	4 5
8. I never worry that people will draw conclusions about my intelligence based on my gender.	1	2	3	4 5
9. Tests, like the one that I just took, have been used to discriminate against people of my gender.	1	2	3	4 5
10. During the test, I wanted to show that people of my gender could perform well on it.	1	2	3	4 5
11. A negative opinion exists about how people of my gender perform on this type of test.	1	2	3	4 5
<i><b>Perceived Performance Pressure</b></i>				
12. I felt a great deal of pressure to perform well on this test.	1	2	3	4 5
<i><b>Motivation</b></i>				
13. Doing well on this test was important to me	1	2	3	4 5
14. I wanted to do well on this test.	1	2	3	4 5
15. I tried my best on this test.	1	2	3	4 5
16. I tried to do the very best I could on this test.	1	2	3	4 5
17. While taking this test, I concentrated and tried to do well.	1	2	3	4 5
18. I wanted to be among the top scorers on this test.	1	2	3	4 5

19. I pushed myself to work hard on this test.	1	2	3	4	5
20. I was extremely motivated to do well on this test.	1	2	3	4	5
<b>Anxiety</b>					
21. I probably didn't do as well as most of the other people who took this test.	1	2	3	4	5
22. I am not good at taking tests.	1	2	3	4	5
23. During a test, I often think about how poorly I am doing.	1	2	3	4	5
24. I usually get very anxious about taking tests.	1	2	3	4	5
25. I usually do pretty well on tests.	1	2	3	4	5
26. I expect to be among the people who score really well on this test.	1	2	3	4	5
27. My test scores don't usually reflect my true abilities.	1	2	3	4	5
28. I very much dislike taking tests of this type.	1	2	3	4	5
29. For this test, I found myself thinking of the consequences of failing.	1	2	3	4	5
30. During a test, I get so nervous I can't do as well as I should have.	1	2	3	4	5
31. I am nervous about how my performance on this test will reflect on my abilities.	1	2	3	4	5
32. I am nervous about how my performance on this test will compare to that of others who take it.	1	2	3	4	5
33. I am nervous about what performance on this test means in terms of my intellectual capabilities.	1	2	3	4	5
34. I am nervous about the test scores reflecting my best effort.	1	2	3	4	5
35. I am nervous about the test scores reflecting my true potential for success.	1	2	3	4	5
36. I am nervous about doing my own personal best on this test.	1	2	3	4	5
<b>Competitiveness</b>					
37. I perform better when I am competing against someone rather than when I am the only one striving for a goal.	1	2	3	4	5
38. I do not feel that winning is important in both work and games.	1	2	3	4	5
39. When I win an award or game it means that I am the best compared to everyone else that was playing. It is only fair that the best person win the game.	1	2	3	4	5
40. In school, I always liked to be the first on finished with a test.	1	2	3	4	5
41. I have always wanted to be better than others.	1	2	3	4	5
42. When nominated for an award, I focus on how much better or worse the other candidates' qualifications are as compared to mine.	1	2	3	4	5
43. I would want an A because that means that I did better than other people.	1	2	3	4	5
44. Because it is important that a winner is decided, I do not like to leave a game unfinished.	1	2	3	4	5

## Math Ability Test

INSTRUCTIONS: You will have 25 minutes to complete the following problems. Please mark your answers in the spaces provided on the answer sheet and use the scratch paper as necessary. You may detach the answer sheet and scratch paper from the packet to facilitate page turns.

1. For a given two-digit positive integer, the tens digit is 5 more than the ones digit. The sum of the digits is 11. Find the integer.
2. A theater sells children's tickets for half the adult ticket price. If 5 adult tickets and 8 children's tickets cost a total of \$27, what is the cost of an adult ticket?
3. John drives from his house to a beach 150 miles away, and at the end of the day drives home. If he drives at an average of 50 miles per hour, how long does the round trip take?
4. If Jane has 5 pairs of pants and 7 shirts, how many different combinations of pants and shirts are possible?
5. Lenny's average score after 3 tests is 88. What score on the 4th test would bring Lenny's average up to exactly 90?
6. A florist buys roses at \$0.50 a piece and sells them for \$1.00 apiece. If there are no other expenses, how many roses must be sold in order to make a profit of \$300?
7. Six cups of flour are required to make a batch of cookies. How many cups of flour are needed to make enough cookies to fill 12 cookie jars, if each cookie jar holds 1.5 batches?
8. A business is owned by 9 women and 1 man, each of whom owns an equal share. If one of the women sells  $\frac{1}{2}$  of her share to the man, and another woman keeps  $\frac{1}{5}$  of her share and sells the rest to the man, what fraction of the business will the man own?
9. If Leah is 6 years older than Sue, and John is 5 years older than Leah, and the total of their ages is 41. Then how old is Sue?

10. If 1 alpha = 2 betas and 1 beta = 3 gammas, how many alphas are equal to 36 gammas?
11. Jake's average test score after 2 tests is 78. What average score must Jake score on the 3rd, 4th, and 5th tests to bring his average up to exactly 87?
12. Grace has 16 jellybeans in her pocket. She has 8 red ones, 4 green ones, and 4 blue ones. What is the minimum number of jellybeans she must take out of her pocket to ensure that she has one of each color?
13. The sales price of a car is \$12,590, which is 20% off the original price. What is the original price?
14. Dan drives to Cheryl's house at an average speed of 40 mph. If he can drive  $\frac{2}{3}$  of the way there in an hour, how far away is Cheryl's house?