


Spring 2014

Gender Stereotypes in Leadership: How Threatening Are They?

Valerie N. Streets
Old Dominion University

Follow this and additional works at: https://digitalcommons.odu.edu/psychology_etds

 Part of the [Experimental Analysis of Behavior Commons](#), [Feminist, Gender, and Sexuality Studies Commons](#), and the [Industrial and Organizational Psychology Commons](#)

Recommended Citation

Streets, Valerie N.. "Gender Stereotypes in Leadership: How Threatening Are They?" (2014). Master of Science (MS), thesis, Psychology, Old Dominion University, DOI: 10.25777/gew0-c787
https://digitalcommons.odu.edu/psychology_etds/108

This Thesis is brought to you for free and open access by the Psychology at ODU Digital Commons. It has been accepted for inclusion in Psychology Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.

GENDER STEREOTYPES IN LEADERSHIP: HOW THREATENING ARE THEY?

by

Valerie N. Streets
B.S., May 2011, Towson University

A Thesis Submitted to the Faculty of Old Dominion University in Partial Fulfillment of
the Requirements for the Degree of

MASTER OF SCIENCE

PSYCHOLOGY

OLD DOMINION UNIVERSITY
May 2014

Approved by:

Debra A. Major (Director)

James P. Bljss (Member)

Ivan K. Ash (Member)

ABSTRACT

GENDER STEREOTYPES IN LEADERSHIP: HOW THREATENING ARE THEY?

Valerie N. Streets
Old Dominion University, 2014
Director: Dr. Debra A. Major

Women's persistent underrepresentation in management has prompted a considerable body of research to better understand how gender stereotyping contributes to this disparity. One possible explanation for the impact of stereotyping on women in management is stereotype threat (i.e., the risk of confirming negative stereotypes as true). Experimental research concerning stereotype threat as it affects women within the domain of leadership has been limited, with no published study specifically manipulating stereotype threat and testing effects on subsequent leadership performance. This thesis expands upon the current literature by replicating classic stereotype threat experimental designs and applying such a design to a leadership context. No significant performance effects or affective reactions to a stereotype trigger were found. Implications, limitations, and future research directions are presented.

ACKNOWLEDGMENTS

I would like to thank Dr. Debra A. Major for mentoring me and continually providing the ideal combination of autonomy and guidance to truly grow as an academic. I would also like to thank my committee members, Dr. Ivan K. Ash, and Dr. James P. Bliss, for their assistance. Special thanks go to my peers for their encouragement, assistance, and commiseration throughout this process. I'm lucky to call you all friends. Lastly, I would like to extend a thank you to my parents for their unconditional love and support, without which I would not be on this career path.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vi
 Chapter	
I. INTRODUCTION	1
WOMEN IN MANAGEMENT	1
GENDER STEREOTYPING	3
WOMEN AND LEADERSHIP	5
STEREOTYPE THREAT	6
WOMEN AND STEREOTYPE THREAT	8
STEREOTYPE THREAT AND LEADERSHIP	9
THE CURRENT STUDY	12
II. METHOD	14
PARTICIPANTS	14
PROCEDURE	15
MEASURES	17
III. RESULTS	19
Data Screening	19
Hypothesis 1	20
Hypothesis 2	22
A Comparison of MTURK and SONA Participants	23
IV. DISCUSSION AND CONCLUSIONS	26
Limitations	30
Future Research Directions	32
REFERENCES	35
 APPENDICES	
A. NOTIFICATION STATEMENT	46
B. MANIPULATION	48
C. LEADER EFFECTIVENESS AND ADAPTABILITY DESCRIPTION (LEAD)	49
D. PERCEIVED STEREOTYPE THREAT	52
E. EXPERIMENTER INSTRUCTIONS AND SCRIPT	53
F. DESCRIPTIVE STATISTICS OF OUTCOME MEASURES	55

VITA.....56

LIST OF TABLES

Table		Page
1.	Analysis of Variance for Leadership Decision Making Scores of MTurk Participants.....	21
2.	Analysis of Variance for Leadership Decision Making Scores of SONA Particiapnts.....	22
3.	Comparison of Mean Perceptions of Stereotype Threat among Women.....	23
4.	Analysis of Covariance for Age and LEAD Performance.....	25
5.	Descriptive Statistics of Outcome Measures by Condition and Data Source.....	55

CHAPTER I

INTRODUCTION

Women's persistent underrepresentation in management (Catalyst, 2012) has prompted a considerable body of research to better understand how gender stereotyping contributes to this disparity (e.g., Heilman, 2001). However, it is imperative to consider the ways in which stereotyping is manifested and the repercussions these manifestations have on women's direct behavioral outcomes (e.g., leadership ability, leadership style) in order to more fully understand the role of stereotyping in this issue. Although gender stereotyping has been evinced as a critical lens through which to view women's underrepresentation, knowledge regarding specific implications of such stereotyping is still lacking. A crucial first step in doing so is to determine whether stereotype threat influences the leadership behavior of women. This research makes such an initial step by exploring the impact of stereotype threat on women in a leadership context.

Women in Management

The dearth of women in management positions has long been covered in the literature (see Wirth, 2009 for a review). It was not until 1996 that women achieved 10 percent representation on corporate boards of *Fortune* 500 companies. Even so, 105 of those companies were lacking any female representation on their boards (Oakley, 2000). Presently, women hold 38.2 percent of U.S. managerial positions (U.S. Department of Labor, 2013), but they are still underrepresented in corporate boardrooms and senior management (Hoobler, Lemmon, & Wayne, 2011). The proportion of U.S. firms with two or more women on their top management teams has yet to exceed 8.5 percent (Dezsö & Ross, 2012).

The presence of women leaders has been demonstrated as a benefit for organizations. First, a diverse team contributes to a good organizational reputation, which corresponds with amicable public relations and better sales (Cox & Blake, 1991). There is also evidence that gender diversity improves innovation levels on management teams. Such diversity provides divergent views and a more comprehensive level of information processing (Van Knippenberg, De Dreu, & Homan, 2004). A survey of Standard & Poor's top 1,500 firms found that firms with women in top management positions perform higher than those lacking such representation; this is especially true of firms with a focus on innovation (Dezsö & Ross, 2012). Additionally, Hoffman and Maier (1961) found that heterogeneous teams (i.e. those with mixed sex and personality composition) yield higher quality problem solving skills. A more gender-inclusive management team is also linked to increased levels of organizational flexibility. This is because the increased tolerance for diversity corresponds to a generally increased level of openness to new ideas (Cox & Blake, 1991).

The above findings have led to an array of proposed explanations for the gender disparity in leadership. The lack of women executives is partially attributed to formal policy matters and organizational practices such as recruitment and promotion tactics. For example, CEO positions require previous line experience within marketing or operations departments, an opportunity few women have. Women occupying middle-level management positions also report receiving less performance feedback relative to their male counterparts, further hindering advancement (Oakley, 2000)

Liff and Ward (2001) studied junior and senior managers in a U.K. bank and found that management, particularly at the senior level, was viewed as a male preserve.

Their findings pointed to informal practices as strong explanations for the gender gap. Top-tier managers were perceived as workaholics who devoted themselves entirely to their work. Many women in the sample reported discomfort with this image and consequential disinterest in such positions. Additionally, women seem to be barred from advancement to the top because they represent a threat to the status quo, or the tradition of an “old boys’ network.” This network is an informal male social network at the top levels of organizations and is exclusionary toward lower-status men and all women (Lipman-Blumen, 1976). As women begin to break into such a network, competency testing is often implemented, where women must repeatedly prove themselves to their male colleagues (Rosener, 1995).

While several explanations for women’s underrepresentation in management have enjoyed empirical support, the gender disparity is still not fully understood. The issue is a complex one, with multiple explanations. This research considers stereotyping, and specifically stereotype threat, as a phenomenon that helps explain the underrepresentation of women in leadership positions.

Gender Stereotyping

Gender stereotypes are largely centered upon women’s and men’s occupation of distinct social roles. Specifically, women are more likely to assume the roles of homemaker and caregiver whereas men are more likely to be breadwinners. From this distinction, role-based differences, or different behavioral and personality traits conferred about the occupants of each role, emerge (Eagly, 1997). These stereotypes exist on an agency-community continuum, with men being perceived as more agentic (i.e., assertive, independent, achievement-oriented) and women as more communal (i.e., warm,

nurturing, compassionate; Eagly & Steffen, 1984).

However, gender stereotypes go beyond the roles held by women and men. Taylor, Fiske, Etcoff, and Ruderman (1978) demonstrated that such stereotyping exists when the roles and behaviors of the target remain constant. Furthermore, these stereotypes apply to adults and young children alike, thereby showing that they are not solely based on the roles occupied by adult women and men (Condry & Condry, 1976). Thus, rather than being based on a sexual division of labor, it is likely that gender stereotypes exist to justify that role allocation. It has proven historically beneficial for men and women to assume different roles; gendered division of labor provided a means of dividing and conquering. Thus, it has become adaptive to believe that each gender is best suited for the requisite performance in each type of role (Williams & Best, 1982). These beliefs about the capabilities and talents of each gender translate to personality traits, as there are widely held beliefs that men and women have inherent differences that predispose them for distinct personalities (Schmitt, Realo, Voracek, & Allik, 2008). Society utilizes these differences in personality and ability to justify the extant discrepancies in the statuses and roles occupied by each gender (Hoffman & Hurst, 1990).

Gender stereotypes exist in two forms. Descriptive stereotyping entails expectations of what women are like or beliefs about women's traits and behaviors. Prescriptive stereotyping is comprised of beliefs regarding the ways in which women should conduct themselves. Prescriptive gender stereotypes assert that men should display agentic traits such as independence, assertiveness, dominance, and competence. In contrast, women should display communal traits such as friendliness, emotional

expressiveness, nurturance, and compassion (Eagly & Wood, 1991). Both types of stereotyping ultimately lead to the devaluation of women's job performance and unfavorable evaluations of female employees (see Eagly, Makhijani, & Klonsky, 1992 for a review).

Women and Leadership

Schein (1975) introduced the think-manager – think-male paradigm, which asserts that traits assumed to be typical of men align much more closely with those of effective leadership than do those that are ascribed to women. This phenomenon continues to receive empirical support (e.g. Heilman, Block, & Martell, 1995; Powell, Butterfield, & Parent, 2002). Employees are generally viewed through a gendered lens in which cultural beliefs about the roles of women and men dictate expectations and perceptions of individuals (Eagly & Karau, 2002). Such stereotypes generate from observations of individuals in gender-typical social roles; thus, communal traits such as warmth and sensitivity are expected of women, while agentic traits such as independence and assertiveness are ascribed to men (Cann & Siegfried, 1990). These stereotypes are deeply rooted within Western culture and are thereby difficult to override (Heilman, Block, Martell, & Simon, 1989).

Despite these stereotypes, research has demonstrated little to no difference in the ways women and men actually lead (e.g. Andersen & Hansson, 2011; Hyde, 2005; Miller Burke & Attridge, 2011), especially when they occupy the same managerial positions, as they have been selected based on the same organizational criteria. However, differences in levels of effectiveness yielded by men and women leaders have been shown, with men being more successful than women in positions defined in masculine terms and women

enjoying more success than men when performing in roles defined in more feminine terms. These effectiveness findings exist despite apparent equality in abilities and qualifications (Eagly, Karau, & Makhijani, 1995). Such findings make a case for the role of stereotype threat for women in the domain of leadership, as negative stereotypes regarding women's capacities are present, and factors other than ability seem to contribute to performance decrements.

Stereotype Threat

Stereotype threat occurs when members of a given sociodemographic group are aware of a negative stereotype regarding the abilities of their group. This awareness limits members' performance within the stereotyped domain (Steele & Aronson, 1995). Cognizance of these stereotypes can be triggered blatantly, with the explicit presentation of stereotypical information, or subtly by mere recognition of one's sociodemographic identity (Hoyt & Blascovich, 2010). Effects have been demonstrated across varying groups and domains, such as: African Americans and Latinos in standardized testing (Gonzales, Blanton, & Williams, 2002; Steele & Aronson, 1995), women in math and science (Shapiro & Williams, 2012), whites in athletics (Stone, Lynch, Sjomeling, & Darley, 1999), the elderly in memory (O'Brien & Hummert, 2006), and men in social sensitivity (Koenig & Eagly, 2005). Recently, stereotype threat has begun to be explored with regard to its implications for women's leadership (e.g., Davies, Spencer, & Steele, 2005; Hoyt, Johnson, Murphy, & Skinnell, 2010).

Stereotype threat effects (i.e., performance decrements) are believed to emerge through a number of operational mechanisms. Threat has been demonstrated to heighten physiological arousal among threatened individuals (Blascovich, Spencer, Quinn, &

Steele, 2001; Osborne, 2006), which has been empirically linked to decreased performance on difficult tasks (O'Brien & Crandall, 2003). Such arousal can reduce self-regulation of attention and behavior, further hindering performance (Inzlicht, McKay, & Aronson, 2006). Cognitive and emotional mechanisms have yielded mixed findings in the research. There exists some suggestion that stereotype threat triggers anxiety prior to performance on a task (Marx & Stapel, 2006). Lowered performance expectations and reduced self-esteem have also been shown to lead to performance decrements in stereotype threatened individuals (Cadinu, Maass, Frigerio, Impagliazzo, & Latinotti, 2003; Keller & Dauenheimer, 2003). Ultimately, such responses, either individually or concomitantly, are believed to hinder performance through the depletion of the cognitive resources one has available to devote to a given task (Johns, Inzlicht, & Schmader, 2008). Awareness of the negative stereotype and its relevance to the given task elicit worrisome thoughts and distract individuals, thereby hindering focus on performance.

The detrimental effects of stereotype threat are most likely to emerge in the presence of specific contextual factors. First, the salience of a stereotyped identity is critical; women and racial minorities are generally more vulnerable in that they have highly visible stereotyped identities (Quinn, Kahng, & Crocker, 2004). This effect can be exacerbated in situations where one anticipates being or is actually in the minority of a group, especially when the individual is the sole representative of a stereotyped group (Murphy, Steele, & Gross, 2007; Sekaquaptewa, Waldman, & Thompson, 2007). Additionally, a negative stereotype must be made salient. This can be done with explicit expression of the stereotype (e.g., Spencer, Steele, & Quinn, 1999), or with the framing of the task; the more closely the description of a task resembles the stereotype, the more

likely threat effects are to emerge. For example, labeling a task as a difficult math test can lead to lower performance than when the same test is presented as a measure of academic self-regard (Yopyk & Prentice, 2005).

Women and Stereotype Threat

Much of the research on gendered implications of stereotype threat focuses on women's performance on math tasks (e.g., Keller, 2002; Oswald & Harvey, 2000; Spencer et al., 1999). Despite the fact that stereotype threat research about women has been housed in one domain, factors unique to women's general experience of stereotype threat have been uncovered. Women are more likely than members of other stereotyped groups to employ behavioral avoidance in response to stereotype threat (Davies, Spencer, Quinn, & Gerhardstein, 2002; Murphy et al., 2007). Additionally, women have demonstrated higher levels of thought suppression concerning negative stereotypes than have other stereotyped groups (Logel, Iserman, Davies, Quinn, & Spencer, 2009).

Although women have an easier time suppressing negatively stereotyped thoughts, they are more inclined than other groups to endorse the stereotypes that apply to their identity (Jackman, 1994). This is likely because, unlike ethnic minorities, women are affected by stereotype threat in ways reflective of both positive and negative aspects of their identity; women are often perceived as incompetent in math and leadership domains, but also as generally warm and compassionate (i.e., likable) people (Logel et al., 2009). Perhaps because of the ambivalent nature of gender stereotyping, women are also less cognizant of the effects of stereotyping and stereotype threat. Women report lower levels of stigma consciousness than men, thereby making them the only

stigmatized sociodemographic group to report less stigma consciousness than a non-stigmatized group (Pinel, 1999).

Stereotype Threat and Leadership

Although research on stereotype threat for women within a leadership context is currently limited, two major themes have emerged: implications for leadership aspirations and possible moderators of the stereotype threat – performance relationship. With regard to stereotype threat and leadership aspirations, Davies et al. (2005) began examining this topic as influenced by media messages. Specifically, they evaluated the ways in which exposure to commercials depicting individuals in gender-stereotyped roles impacted participants' preferences for assuming leadership roles. They found that women exposed to gender-stereotyped commercials had a greater aversion to leadership roles than women in a gender-neutral commercial condition. Moreover, women in the gender-stereotyped condition avoided leadership tasks as a means of avoiding confirmation of stereotypes. Gupta, Turban, and Bhawe (2008) extended these findings by assessing the effects of explicit and implicit gender stereotype activation on intentions to pursue a specific avenue of leadership (i.e., entrepreneurship). Participants read an article describing characteristics of effective entrepreneurs in terms of either masculine or feminine traits. Participants in explicit conditions were also told that entrepreneurs display characteristics typical of American masculinity or femininity. When the role was implicitly linked to masculinity, women were much more likely to report entrepreneurial intentions than when the role was blatantly presented as masculine. Such a finding implies that more blatant stereotype activation triggers greater levels of stereotype threat, thereby inhibiting women's leadership aspirations. Proactive personality (i.e., a stable

disposition toward resisting situational constraints and affecting change; Bateman & Crant, 1993) has also been proposed as a moderator of the relationship between stereotype threat and women's leadership aspirations. Entrepreneurial intentions of women with higher levels of proactive personality are less affected by stereotype threat than those of women with lower levels of the trait (Gupta & Bhawe, 2007).

Regarding moderators of threat effects, leadership self-efficacy is the most explored. Leadership self-efficacy has largely been demonstrated as a buffer against stereotype threat effects. Hoyt (2005) found that women high in leadership self-efficacy were more likely to identify strongly with the leadership domain following stereotype activation than were less efficacious women, suggesting self-efficacy as another buffer against potentially detrimental effects of stereotype threat. Additionally, high levels of leadership self-efficacy prior to stereotype threat activation ameliorate the decrements to women's self-esteem and leadership self-efficacy following stereotype triggers (Burnette, Pollack, & Hoyt, 2010). The moderating role of leadership-self efficacy has also been explored with regard to its impact on actual leadership performance. Parallel to the findings of Hoyt (2005) and Burnette et al. (2010), Hoyt and Blascovich (2007, 2010) found that highly efficacious women were less vulnerable to gender stereotype activation as indicated by both behavioral and self-report measures (i.e., perceived performance, domain identification, and psychological well-being).

Bergeron, Block, and Echtenkamp (2006) explored gender identification as another possible moderator of stereotype threat and leadership performance. Contrary to other previous research (e.g., Hoyt & Blascovich, 2007; Hoyt et al., 2010), the authors did not manipulate stereotype threat, but rather measured levels of stereotype threat

experienced within the workplace by using a field sample. A main effect for gender was revealed, with women experiencing more stereotype threat within the domain of leadership than men.

As outlined, the extant research within the leadership domain focuses either on threat effects on leadership aspirations or potential moderators of the threat-performance relationship. However, researchers have not manipulated stereotype threat and explicitly measured subsequent leadership performance. Rather, threat effects have been assumed on the basis of demonstrated decrements in outcomes such as leadership aspirations and perceived performance. Thus, this research seeks to demonstrate that stereotype threat inhibits the actual leadership performance of women. Based on the patterns demonstrated with self-efficacy, leadership aspirations, domain identification, and self-reported performance, it is posited that stereotype threat activation will lead to performance decrements for women within the domain of leadership.

Hypothesis 1: Women exposed to a stereotype threat manipulation will demonstrate lower levels of leadership performance than will women in the control condition and men in either the experimental or control condition.

Furthermore, the current research will evaluate participants' levels of perceived stereotype threat. This will identify whether threatened women are cognizant of the impact of gender stereotypes. The findings that stereotype threat activation elicits consciously motivated outcomes such as a reduced willingness to lead (e.g., Davies et al., 2002) suggest that stereotype threat does act within women's awareness. Thus, it is hypothesized that women in the stereotype threat condition will report higher levels of perceived stereotype threat. Because men are not stereotype threatened within the

leadership context, they will not be evaluated on this construct.

Hypothesis 2: Women exposed to a stereotype threat activation will report higher levels of perceived stereotype threat than will women in the control condition.

The Current Study

The extant research on stereotype threat and leadership is predicated on a considerable inferential leap: stereotype threat is presumed to occur for women within the leadership domain based on proxies for leadership behavior. Previous studies have sought to uncover stereotype threat within this realm by assessing constructs such as leadership self-efficacy (e.g., Burnette et al., 2010) or willingness to lead (e.g., Davies et al., 2005) as outcome variables. However, research within other domains has uncovered stereotype threat by revealing a true performance decrement (e.g., Steele & Aronson, 1995). The only leadership study to measure performance as a dependent variable was that of Bergeron et al. (2006). However, the lack of an experimental manipulation prevents a causal inference from being drawn. Thus, this study combines the experimental control of previous studies of leadership performance proxies (e.g. Hoyt, 2005) with the measurement of leadership performance initiated by Bergeron et al. (2006).

Stereotype threat research in other domains generally began with a study demonstrating the existence of stereotype threat via performance on a standardized test. For example, Steele and Aronson (1995) introduced stereotype threat by demonstrating that the advertised diagnostic ability of a standardized verbal test led to lower performance among Black participants compared to White participants. Additionally,

Spencer et al. (1999) extended this threat to women and mathematics by showing that describing a test as illustrative of gender differences yielded worse test performance for women relative to men. Regarding women and stereotype threat, subtle manipulations (e.g., emphasizing the diagnostic ability of a test) have evoked larger stereotype threat effects than have blatant manipulations (e.g., stating the group's inferiority on the given task; Nguyen & Ryan, 2008). Using previous stereotype threat research as a guide, the current study used a subtle manipulation to test the existence of stereotype threat effects on women's leadership performance. By demonstrating an effect for women on a test of leadership ability, this study attempted to uncover the existence of a true stereotype threat effect for women performing in a leadership context.

CHAPTER II

METHOD

Participants

A power analysis was conducted with G-Power to determine an appropriate sample size for this research. Small effect sizes have been revealed in previous research linking stereotype threat to women's performance outcomes (Nguyen & Ryan, 2008). The power analysis used a small-to-medium effect size ($d = .43$) for a t-test comparing independent groups. Results indicated that 68 participants will be needed in each group. Therefore, a total sample of 272 will be necessary to detect similar effects in the current study.

Data were collected in two waves. Amazon's Mechanical Turk (MTurk) was used to collect the first round of data. Mechanical Turk was launched in 2005 and is a means of crowd-sourcing labor intensive tasks. It has recently been adopted as a source of research participants in Psychology, due to its availability of a large and diverse subject pool, low costs, and brief turnaround times (Crump, McDonell, & Gureckis, 2013; Mason & Suri, 2012). Some differences between MTurk and laboratory participants have been identified: MTurk participants are less likely to pay attention to experimental manipulations, are more likely to research answers on the Internet, and report lower levels of extraversion and self-esteem than laboratory participants (Goodman, Cryder, & Cheema, 2012). Given the type of task in the current study and the brief experimental manipulation, these differences were not viewed as problematic for the current study. Furthermore, MTurk has been demonstrated as a viable vehicle for collecting data in judgment and decision making research (Paolacci, Chandler, &

Ipeirotis, 2010). Because participants were asked to complete a situational judgment test, such findings were supportive of using MTurk as a data collection medium.

Three hundred and six individuals completed the study on Mechanical Turk. Sixty-two participants were eliminated from analyses due to failure to provide information regarding gender ($n = 11$), failure to complete the manipulation check ($n = 1$), and failure to complete the leadership test in the allotted ten minutes ($n = 50$). This resulted in a final sample size of 244. Of this sample, 57.8 percent were female ($n = 141$) and the average age was 36.93 ($SD = 12.88$). The majority of the sample were employed ($n = 179$) with fields such as customer service, healthcare, human resources, and retail represented. Sixteen percent of the sample were students ($n = 39$), with most of those participants being graduate or professional students ($n = 14$).

The second wave of data was collected in person via the SONA participant pool. Students were awarded one credit for completing the study at a designated time in a computer lab. Two hundred participants completed the study. Of that sample, 39 individuals were removed from analyses due to failure to provide their gender ($n = 3$), failure to complete the manipulation check ($n = 1$), or failure to complete the leadership test in the allotted ten minutes ($n = 35$). This resulted in a final sample of 161. Of this sample, 73.3 percent were female ($n = 118$) and the average age was 20.43 ($SD = 5.58$).

Procedure

Participants were grouped following a 2 (male, female) x 2 (threat, non-threat) design. Men and women were randomly assigned to either a stereotype threat or a non-threat condition. The first wave of data was collected via Mechanical Turk. Participants

had a maximum of one hour to complete the task. Once participants opted to participate in the study, they were provided with a notification statement (Appendix A). Participants were then asked to provide their Amazon Mechanical Turk worker identification number to receive compensation for their participation. Participants were asked to complete the tasks in a setting that was free from distractions. Once they were in such a location they were prompted to read the instructions on the next screen (Appendix B). Here the manipulation, which has been adapted from Spencer et al. (1999), was introduced. The screen informed participants in the control condition that ODU is developing a series of new tests to use in future research and they would be administered one at random. Participants in the stereotype threat condition read a message informing them that the research was testing their leadership ability.

Following the introduction, participants were instructed to complete a leadership situational judgment test designed by Hersey and Blanchard (1976; Appendix C). The measure consists of 12 scenarios for which respondents are asked to select the best course of action. Following completion of the situational judgment test, participants were administered a measure of perceived stereotype threat (McIntyre, Paulson, Taylor, Morin, & Lord, 2011; Appendix D). Because men should not be stereotype threatened by a leadership task, the questions on the original measure are not as applicable to men in this context (e.g., “I worried that the researcher will think that women as a whole have less leadership ability because of how I did on this test”). Thus, the measure was adapted for male participants. Lastly, participants were asked to report demographic characteristics, including gender. This question was reserved until the end of the session so as to avoid introducing another potential stereotype threat trigger.

The second wave of data was collected in person with participants from the SONA research pool. Both participants and experimenters were blind to the condition. Studies of stereotype threat affecting women within mathematics have often run mixed-gender groups (e.g., Brown & Josephs, 1999; Davies et al., 2002; Martens, Johns, Greenberg, & Schimel, 2006). Thus, the current design employed mixed-gender groups in the computer lab (Mills Godwin Building, Room 222). Upon arrival, participants were seated at individual computers. Participants were greeted by a female experimenter and informed that the researchers were developing new tests to be evaluated across a large group of Old Dominion University students. Experimenters were provided with a script (Appendix E) which they recited to all participants at the beginning of the session. Then participants were provided with the same notification statement, instructions, and series of measures as were Mechanical Turk participants. As with the first round of data collection, participants were given a maximum of one hour to complete the task.

Measures

Leadership performance. Leadership performance was assessed with the leadership situational judgment test Leader Effectiveness and Adaptability Description (LEAD; Hersey & Blanchard, 1976; Appendix C). The test was designed to be completed in 10 minutes and contains 12 scenarios depicting general circumstances encountered by leaders. For each scenario, participants are asked to choose the best of four response options. Each item is designed to represent one of four styles of leadership: telling (high task-orientation, low relationship-orientation), selling (high task-orientation and relationship-orientation), participating (low task-orientation, high relationship-orientation), and delegating (low task-orientation and relationship-orientation). While the

instrument can be used to determine the primary style with which an individual leads, it also asserts that leaders should be able to adapt their style according to the demands of the situation. Thus, item scores were also summed to calculate an overall leadership effectiveness score, which served as the dependent variable. Scores can range from 0 to 36, with higher scores indicating greater leadership effectiveness. The instrument has been standardized on the basis of responses from 264 business managers and had a reported coefficient alpha of 0.71 (Greene, 1980). However, substantially lower alpha values were obtained in the current study (i.e., 0.30 and 0.45 for mTurk and SONA data, respectively).

Perceived stereotype threat. Participants' sense of experienced stereotype threat was assessed with a six-item scale adapted from McIntyre et al. (2011; Appendix D). The scale, originally designed for a sample of stereotype threatened women with regard to math performance, was adapted to reflect the leadership context of the current study. Participants responded on a scale from 1 (*not at all*) to 7 (*very much*) to questions such as, "how often did you think about performing poorly while you took the leadership test?" The measure had a reported coefficient alpha of 0.77. In the current study, a coefficient alpha of 0.85 and 0.82 were obtained for mTurk and SONA data respectively.

CHAPTER III

RESULTS

This chapter describes and summarizes the statistical analyses used to evaluate the hypotheses established in previous chapters. Subsequent to the data screening process, this chapter reports the results of the analyses of variance conducted to evaluate Hypothesis 1. An evaluation of Hypothesis 2 on the basis of independent-samples *t* tests is also reported.

Data Screening

All data points were screened to ensure they met three preliminary criteria. First, because the Leader Effectiveness and Adaptability Description was designed to be completed within ten minutes and participants were notified as such, any individuals who took longer than ten minutes to complete the test were excluded from the analyses. Second, the necessary analyses are predicated upon knowing the gender of all participants; thus, any participants who did not provide their gender were excluded from all analyses. Lastly, all participants were provided with a manipulation check in which they were asked to retype the instructions they viewed. Because the experimenter was blind to participants' conditions, entry of the instructions was necessary to confirm condition status. Failure to provide this information also resulted in exclusion from analyses.

Once preliminary screening was completed, the data were screened for outliers and checked for the necessary assumptions of analysis of variance. First, outliers on the basis of LEAD completion time were checked. Standardized leadership scores were created and the data set was checked for any scores with an absolute value greater than or

equal to three. No Mechanical Turk or SONA participants were considered outliers by this criterion. Atypical response behavior was gauged by further screening test completion time for outliers. As with leadership, a standardized timing variable was created. No individuals had a standardized test time variable with an absolute value of three or greater.

To check the first assumption of analysis of variance, data were screened for normality and skewness and kurtosis values for LEAD scores were acceptable (i.e., less than the absolute value of two). The assumption of homogeneity of variance was assessed via a Levene's test. For the Mechanical Turk, $F(1, 136) = 2.391, p = .142$ and SONA data $F(3, 157) = 2.447, p = .066$, test results were not significant, confirming that this assumption was met. Lastly, independence of observations was satisfied by random assignment of participants to threat and non-threat conditions. Descriptive statistics are presented in Appendix F.

Hypothesis 1

Hypothesis one was tested with a 2 x 2 analysis of variance and followed up with a planned contrast, such that test scores of women in the stereotype threat condition were compared to those of participants in the other three groups. Because gender stereotypes exist in such a way that they are detrimental to women, no difference was expected between the scores of men in the stereotype threat condition and those of men in the control condition. Furthermore, in the absence of apparent stereotyping or sexism, a significant difference in the effectiveness of male and female leadership has not been demonstrated (see Eagly et al., 1995 for a review). Thus, no differences between women in the control condition and men in either condition were expected.

For the Mechanical Turk data, a 2 x 2 ANOVA revealed no significant main or interaction effects (Table 1). The non-significant main effects for gender and condition reveal that leadership scores did not differ on the basis of gender or exposure to the manipulation. Thus, performance in each of the four groups does not meaningfully differ from that of any other group and hypothesis one was not supported. To ensure that no stereotype threat effect was elicited, an a priori contrast was conducted and also revealed no difference between the performance of women who received the manipulation and that of participants in any other group, $F(1, 240) = 0.087$ $p = .768$. Group means and standard deviations are presented in Appendix F.

Table 1.
Analysis of Variance for Leadership Decision Making Scores of MTurk Participants

Source	<i>df</i>	<i>MS</i>	<i>F</i>	Significance	Partial η^2
Gender	1	20.185	1.496	.222	.006
Condition	1	5.146	0.381	.537	.002
Gender*Condition	1	0.163	0.012	.913	.000
Error	240	13.493			

Laboratory data were analyzed using a modification of the above procedure. The available participant pool was female dominated, making it difficult to secure sufficient male participation in a timely manner. After one month of data collection, once participation rates began to slow, an independent samples *t*-test was conducted to determine whether test performance differed for women on the basis of assigned condition. At the time of preliminary analysis, 43 men had participated and 118 women had participated. The *t*-test revealed no difference between the two groups, $t(118) = 0.241$, $p = .810$. Because the necessary number of women had participated in the study

and no effects from the manipulation were found, there was justifiable means to terminate data collection.

As with the Mechanical Turk data, the laboratory data were analyzed via a 2 x 2 ANOVA, however the incongruent group sizes and absence of a manipulation effect prompt caution in any interpretation of the following results. A two-way ANOVA indicates a significant difference between men and women on the leadership test, $F(1, 157) = 180.709, p < .001, \eta^2 = .535$, suggesting that men demonstrated significantly higher scores on the test when compared to women. However, the ANOVA results fail to indicate that there is a main effect for condition or an interaction effect between gender and condition (Table 2). Group means and standard deviations are also presented in Appendix F. Because the three groups to be compared to women who received the manipulation (i.e., men in the threat condition, men in the non-threat condition, and women in the non-threat condition) were not equivalent, the proposed contrast does not make sense to conduct. Thus, it can be concluded that Hypothesis 1 was not supported.

Table 2.

Analysis of Variance for Leadership Decision Making Scores of SONA Participants

Source	<i>df</i>	<i>MS</i>	<i>F</i>	Significance	Partial η^2
Gender	1	1587.333	180.709	.000	.535
Condition	1	8.406	0.957	.329	.006
Gender*Condition	1	12.715	1.448	.231	.009
Error	157	8.784			

Hypothesis 2

Unlike the first hypothesis, Hypothesis 2 focused on affective reactions to efforts to evoke stereotype threat. Specifically, Hypothesis 2 proposed that women exposed to

the manipulation would report higher levels of perceived stereotype threat than would women who were not exposed to the manipulation. Table 3 summarizes the results of independent-samples t tests conducted to test Hypothesis 2. Among Mechanical Turk participants, women in the threat condition reported no significant difference in perceptions of stereotype threat ($M = 1.688$, $SD = 0.881$) from those in the non-threat condition ($M = 1.475$, $SD = 0.811$), $t(138) = 1.476$, $p = .142$. Similarly, among SONA participants, women in the threat condition demonstrated no significant difference in reports of perceived stereotype threat ($M = 2.214$, $SD = 1.042$) from those in the non-threat condition ($M = 2.409$, $SD = 1.275$), $t(108) = -0.880$, $p = .386$. Thus, both waves of data collection failed to support Hypothesis 2.

Table 3.
Comparison of Mean Perceptions of Stereotype Threat among Women

Data Source	Threat Group		Non-Threat Group		t	Sig.
	M	SD	M	SD		
MTurk	1.688	0.881	1.475	0.811	1.476	.142
SONA	2.214	1.042	2.409	1.275	-0.880	.386

A Comparison of MTurk and SONA Participants

Although neither hypothesis was supported in either sample, it is worth noting that SONA and MTurk participants differed in both their leadership test scores and rates of perceived stereotype threat. Specifically, among all participants, those in the MTurk sample received higher test scores ($M = 22.81$, $SD = 3.67$) than did those in the SONA sample ($M = 17.13$, $SD = 4.50$). An independent samples t -test revealed this difference as significant, $t(403) = 13.92$, $p < .001$. Furthermore, MTurk participants reported lower levels of perceived stereotype threat ($M = 1.48$, $SD = 0.81$) than did SONA participants

($M = 2.41$, $SD = 1.28$). This difference was also significant, $t(388) = -6.44$, $p < .001$.

More specifically, women who participated via MTurk reported significantly lower levels of perceived threat than did those who participated via SONA, $t(246) = 5.76$, $p < .001$.

These differences exist despite there being no significant difference in study completion time, missing data, or pass rates for manipulation checks.

An analysis of covariance (ANCOVA) was conducted to further explore the difference on leadership decision making, as this is where a significant gender difference occurred for SONA participants. Specifically, a 2x2x2 ANCOVA was run with gender (male or female), condition (threat or non-threat), and source of data (MTurk or SONA) as the independent variables, age as a covariate, and LEAD performance as a dependent variable. Results of the ANCOVA appear in Table 4. Controlling for age, a significant main effect for gender remained, $F(1,394) = 74.71$, $p < .001$, $\eta^2 = .159$. The persistence of this effect suggests that the underperformance of women in the SONA sample relative to those in the MTurk sample is not due to age differences. Additionally, a significant main effect for data source was revealed, $F(1, 394) = 52.44$, $p < .001$, $\eta^2 = .117$. Furthermore, a significant interaction effect among gender and data source was found, $F(1, 394) = 95.82$, $p < .001$, $\eta^2 = .196$.

Table 4.
Analysis of Covariance for Age and LEAD Performance

Source	<i>df</i>	<i>MS</i>	<i>F</i>	Significance	Partial η^2
Gender	1	860.67	74.71	.000	.159
Condition	1	5.38	0.47	.495	.001
Source	1	604.08	52.44	.000	.117
Gender*Condition	1	14.43	1.25	.264	.003
Gender*Source	1	1103.79	95.82	.000	.196
Condition*Source	1	14.64	1.27	.260	.003
Gender*Condition *Source	1	11.26	0.98	.323	.002
Error	394	11.52			

The significant interaction of gender and source was followed up with a test of simple effects. Controlling for age, the simple effect of gender was significant among SONA participants, $F(1, 394) = 124.67, p < .001, \eta^2 = .240$, but was not significant in the MTurk sample, $F(1, 394) = 1.14, p = .286, \eta^2 = .003$. Among SONA participants, this difference occurred such that women underperformed relative to men. Furthermore, when age was controlled for, the simple effect of data source was significant for women such that SONA women performed worse than did MTurk women, $F(1, 394) = 195.33, p < .001, \eta^2 = .331$, but was not significant for men, $F(1, 394) = 0.52, p = .470, \eta^2 = .001$. These findings suggest that although the experimental manipulation did not elicit stereotype threat, the laboratory environment did produce threat-related performance decrements for women.

CHAPTER IV

DISCUSSION AND CONCLUSIONS

The current study investigated the existence of a stereotype threat effect for women within the context of leadership. It examined the experience of an effect with regard to both performance and affective reactions. In investigating hypothesized stereotype threat effects, no performance effects or perceived stereotype threat effects were found. It is impossible to conclude from a single study that stereotype threat does not exist for women in leadership. Rather, there are a number of plausible explanations for these results.

One such explanation takes into account the differences between waves of data collection. While Hypothesis 1 was not supported for MTurk or SONA participants, a significant main effect for gender emerged among SONA participants. This suggests that stereotype threat may not have been triggered by the manipulation (i.e., the test diagnosticity), but rather was activated by the presence of male participants. In the laboratory, participants completed the study in mixed gender groups. Given that the presence of non-stereotyped others has been demonstrated as sufficient means for eliciting threat effects in other domains (e.g., Stone & McWhinnie, 2008), it is possible that the presence of male test takers primed the negative stereotype for female participants. This is an environmental characteristic that was not present for MTurk participants, and thus could explain the nonsignificant gender effect in the first wave of data collection. Additional research is needed to further explore this postulation, but it is possible that stereotype threat effects do manifest within the domain of leadership.

However, it may be the case that stereotype threat does not hold for women in leadership. An important prerequisite of stereotype threat is the awareness of a negative stereotype regarding one's group (Steele & Aronson, 1995). While women are underrepresented in management, it is possible that their increased presence in managerial occupations is attenuating gender stereotypes. Although women executives are still quite rare, the fact that women hold 38.2 percent of managerial positions means that the image of a female supervisor is likely to be accessible by participants. Thus, leadership may not be relegated as a male domain.

The impact of gender on leadership was widely explored in the 1990s. Research examined the ways in which gender affects leadership styles, leader effectiveness, and evaluations of leadership (Eagly & Johnson, 1990; Eagly et al., 1995; Eagly et al., 1992). However, a critical review of gender stereotypes as they affect leadership has not been conducted in over a decade (Heilman, 2001). Thus, an accurate sense of current gender stereotypes regarding leadership is lacking. Without an updated assessment of societal views toward women and leadership it is impossible to know if negative stereotypes about women in leadership are relevant for the population.

Furthermore, evidence of a female leadership advantage has emerged, in which women are increasingly noted as having leadership styles that promote effective performance (Eagly, 2007; Eagly & Carli, 2003). Recent research has found that women, particularly those that subscribe to more feminine gender identities, demonstrate higher levels of emotional intelligence and transformational leadership (i.e., inspiring and motivating followers to high levels of performance (Lopez-Zafra, Garcia-Retamero, & Pilar Berrios Martos, 2012)). Such evidence suggests that not only may gender

stereotypes about leadership be less relevant now, but beliefs may be shifting in the opposite direction. If this is the case, stereotype threat would not be elicited for women in this domain, as they do not feel negatively stereotyped. Given that no differences in perceived stereotype threat were found on the basis of condition, it is plausible that female participants did not perceive a negative stereotype.

It is also possible that effects were not found because leadership is not an accessible domain for a student sample. Leadership experience is likely more limited among students than it would be in a work sample. Even students who have work experience may not have much experience in a supervisory or leadership role. Given that 49.4 percent of the sample consisted of students ($n = 200$), this sample may not have been an appropriate one in which to assess stereotype threat for leadership.

Lastly, stereotype threat has become controversial, particularly as a topic of organizational research. Some researchers are calling its relevance to applied research into question (Kalokerinos, von Hippel, & Zacher, in press). The lack of extant research analyzing the role of stereotype threat in organizational settings has fostered the view that stereotype threat may be a laboratory phenomenon. While the current research was conducted in a laboratory setting, it differs from the bulk of lab-based stereotype threat research in that an outcome with a clear and objectively correct response was not used. Unlike previous stereotype threat studies, which measured performance on outcomes such as math and general intelligence tests (e.g., Martens et al., 2006), the leadership test in this study was more behavior-oriented. That is, this measure asked participants how they would normally behave in a given situation rather than asking them to actually

perform on a task. If stereotype threat is truly a laboratory phenomenon, measures of this nature are unlikely to be affected by triggers.

Another possible contribution to the absence of an effect is the structure of the outcome measure. Situational judgment tests are often used in selection contexts, particularly for the selection of managers (McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001). However, personnel researchers argue that stereotype threat effects are not relevant in selection contexts because applicant motivation to perform well is high (Sackett, Borneman, & Connelly, 2008; Sackett & Lievens, 2008). Furthermore, several attempts to identify threat effects in selection simulations were not successful (McFarland, Lev-Arey, & Ziegert, 2003; Ployhart, Ziegert, & McFarland, 2003). Given these findings, it is possible that leadership, particularly a situational judgment test designed to measure leadership effectiveness, is not an appropriate criterion for assessing stereotype threat.

Stereotype threat in the purest sense (i.e., a performance decrement suffered because of awareness of a negative stereotype) may not exist within the domain of leadership. The hypotheses of the current research stemmed from findings in other studies that applied stereotype threat to leadership. These studies captured outcomes such as leadership self-efficacy and willingness to assume a leadership role (Burnette et al., 2010; Davies et al., 2002). A true stereotype threat effect was never identified; rather affective reactions to stereotypes were examined. Existence of negative affective outcomes does not necessitate the existence of stereotype threat, and in this case that threat may not exist.

Limitations

One major limitation of this research is that the demonstrated effect sizes in stereotype threat research are very small. Effect sizes among women are particularly small; the mean effect size among stereotype threatened women is a Cohen's d of .205 (Nguyen & Ryan, 2008). Because this study employed a subtle threat manipulation, mean effect sizes for women exposed to subtle threat triggers served as the referent for the current research. Due to the low feasibility of obtaining the necessary sample of male participants via the SONA system, the lower bound of a 95 percent confidence interval ($d = .43$) was used in a priori power analyses. Thus, the sample size necessary to achieve statistical significance may not have been achieved in the current study. A larger sample may have provided the power sufficient to detect significant stereotype threat effects. However, the poor reliability of the measure used to identify threat effects renders this issue irrelevant.

The Leader Effectiveness and Adaptability Description yielded exceptionally low levels of internal consistency. Item analysis failed to reveal any particularly troublesome items, rather the measure as a whole was problematic. Alpha coefficients of this size (0.30 and 0.45) prohibit any solid conclusions from being drawn from the data. Even if hypothesis testing yielded significant results, any interpretation would be ill-advised, as there is no definitive answer to the question "what is being measured?" For example, the coefficient alpha of 0.30, which was obtained from SONA data, indicates that only 30 percent of the variance on the measure is attributable to true score variance (Cortina, 1993). Thus, most of what was captured by the LEAD measure is variance that cannot be accounted for.

Participants were instructed to complete the leadership test while thinking of a situation in which they are most often a leader. Example situations were provided such as a parent or head of the household, a manager at work, or an officer of a student organization. It is possible that allowing for a broad choice of situations introduced too much variability in responses. Although the instructions of the original measure provided the examples of manager and parent, the measure had been validated with managerial samples (Greene, 1980), thereby providing a common referent for the participants. While the scenarios presented on the test were broad, they were not necessarily equally applicable to different types of leadership positions. Consider the sample item, “You stepped into an efficiently run situation. The previous administrator ran a tight ship. You want to maintain a productive situation, but would like to begin humanizing the environment.” This scenario may be easily translated to the workplace but may be less applicable if one’s leadership referent is parenthood.

Furthermore, the choice of a test as an outcome measure may not have been a wise one. While multiple choice tests are commonly used in research to capture the existence of a stereotype threat effect, such tests have not been utilized by leadership researchers. Performance observations and objective performance indicators (e.g., sales or productivity levels) are commonly used as leadership criteria (Hiller, DeChurch, Murase, & Doty, 2011). Following this precedent and capturing actual leadership behavior would have been less abstract of a task for participants and would have been more generalizable to organizational settings.

Future Research Directions

It is difficult to speak to any practical implications of the current research, as the absence of a stereotype threat effect cannot be concluded from this study. However, the limitations and possible explanations for the lack of significant findings create numerous avenues for future research on the topic. In fact, such research is critical in drawing meaningful conclusions from the results of the current study.

One important area of future research is that of gender stereotypes in leadership. An update on the extant literature and a current appraisal of gender stereotypes within this domain is needed to determine if further exploration of stereotype threat in leadership is viable. Specifically, future research is needed to assess whether negative stereotypes about women's ability to lead still exist. If such stereotypes are still present, it is also important to gauge awareness of them. A key contributor to the experience of stereotype threat is the awareness of a negative stereotype (Steele & Aronson, 1995). If negative stereotypes have persisted over time but are no longer as deeply rooted in society, it is unlikely that the stereotype is salient enough to trigger threat.

If gender stereotypes do still permeate the domain of leadership, replication of the current study is needed. Assessment of stereotype threat effects should be done with leadership performance as the criterion. While this is commonly captured via company performance indicators and ratings in the leadership literature, such information would require a field sample. However, an important first step in identifying a stereotype threat effect is laboratory research. Experimental control will be needed to isolate the effect before extending the effect into organizational contexts. Thus, a leadership task that

employs confederates and trained raters may be the best option for future efforts to measure stereotype threat among women in leadership.

Lastly, further investigation of the viability of MTurk as a medium for data collection is needed. MTurk and other such forms of online data collection may mask the presence of certain phenomena. The current research did not find differences in quality, or usability of the data between SONA and MTurk participants. However, some findings suggest that MTurk may not be a proper substitute for laboratory data collection. Scores on both outcome measures differed significantly across waves of data collection. Specifically, a main effect for gender on test performance was demonstrated among SONA participants but not among MTurk participants. If this difference is due to the additional trigger presented by completing the test in person, then it can be concluded that online data collection was not an appropriate medium in which to test this phenomenon.

Additionally, while no gender or condition differences in levels of perceived stereotype threat were revealed in either wave of data collection, overall reports of threat perceptions were higher among SONA participants. Heightened feelings of risk or perceptions of diagnosticity may have been experienced by college students taking the leadership test on a college campus, thereby explaining the increased perceptions of threat by men and women in the SONA sample. Because MTurk participants are not in a test-taking environment, they may not have felt as threatened by the task. These findings suggest that there may be more differences between MTurk and laboratory participants than what has initially been reported in the literature (Goodman et al., 2012). Further

research is needed to detail the differences between MTurk and laboratory research and to determine which types of studies are most suitable for MTurk samples.

REFERENCES

- Andersen, J. A., & Hansson, P. H. (2011). At the end of the road? On differences between women and men in leadership behaviour. *Leadership & Organization Development Journal*, 32(5), 428-441.
- Bateman, T. S., & Crant, J. M. (1993). The proactive component of organizational behavior: A measure and correlates. *Journal of organizational behavior*, 14(2), 103-118.
- Bergeron, D. M., Block, C. J., & Echtenkamp, A. (2006). Disabling the able: Stereotype threat and women's work performance. *Human Performance*, 19(2), 133-158.
- Blascovich, J., Spencer, S. J., Quinn, D., & Steele, C. (2001). African Americans and high blood pressure: The role of stereotype threat. *Psychological Science*, 12(3), 225-229.
- Brown, R. P., & Josephs, R. A. (1999). A burden of proof: Stereotype relevance and gender differences in math performance. *Journal of personality and Social Psychology*, 76(2), 246.
- Burnette, J. L., Pollack, J. M., & Hoyt, C. L. (2010). Individual differences in implicit theories of leadership ability and self-efficacy: Predicting responses to stereotype threat. *Journal of leadership studies*, 3(4), 46-56.
- 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(2), 267-285.
- Cann, A., & Siegfried, W. D. (1990). Gender stereotypes and dimensions of effective leader behavior. *Sex roles*, 23(7-8), 413-419.

- Catalyst. (2012). Women CEOs of the Fortune 1000. Retrieved May 21, 2012, from <http://www.catalyst.org/publication/271/women-ceos-of-the-fortune-1000>
- Condry, J., & Condry, S. (1976). Sex differences: A study of the eye of the beholder. *Child development, 47*, 812-819.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98-104.
- Cox, T. H., & Blake, S. (1991). Managing cultural diversity: Implications for organizational competitiveness. *The Executive, 5*(3), 45-56.
- Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS ONE, 8*(3): e57410. doi:10.1371/journal.pone.0057410
- Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin, 28*, 1615-1628.
- Davies, P. G., Spencer, S. J., & Steele, C. M. (2005). Clearing the air: Identity safety moderates the effects of stereotype threat on women's leadership aspirations. *Journal of personality and Social Psychology, 88*(2), 276-287.
- Dezsö, C. L., & Ross, D. G. (2012). Does female representation in top management improve firm performance? A panel data investigation. *Strategic Management Journal, 33*, 1072-1089.
- Eagly, A. H. (1997). Sex differences in social behavior: Comparing social role theory and evolutionary psychology. *American Psychologist, 52*, 1380-1383.

- Eagly, A. H. (2007). Female leadership advantage and disadvantage: resolving the contradictions. *Psychology of Women Quarterly, 31*, 1-12.
- Eagly, A. H., & Carli, L. L. (2003). The female leadership advantage: An evaluation of the evidence. *The Leadership Quarterly, 14*, 807-834.
- Eagly, A. H., & Johnson, B. T. (1990). Gender and leadership style: A meta-analysis. *Psychological Bulletin, 108*, 233-256.
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109*, 573-598.
- Eagly, A. H., Karau, S. J., & Makhijani, M. G. (1995). Gender and the effectiveness of leaders: A meta-analysis. *Psychological Bulletin, 117*, 125-145.
- Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin, 111*, 3-22.
- Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology, 46*, 735-754.
- Eagly, A. H., & Wood, W. (1991). Explaining sex differences in social behavior: A meta-analytic perspective. *Personality and Social Psychology Bulletin, 17*, 306-315.
- Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of Latino women. *Personality and Social Psychology Bulletin, 28*, 659-670.
- Goodman, J. K., Cryder, C. E., & Cheema, A. (2012). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making, 26*, 213-224.

- Greene, J. F. (1980). *A summary of technical information about LEAD-Self*. San Diego: Learning Resources Corporation.
- Gupta, V. K., & Bhawe, N. M. (2007). The influence of proactive personality and stereotype threat on women's entrepreneurial intentions. *Journal of Leadership & Organizational Studies, 13*(4), 73-85.
- Gupta, V. K., Turban, D. B., & Bhawe, N. M. (2008). The effect of gender stereotype activation on entrepreneurial intentions. *Journal of Applied Psychology, 93*, 1053-1061.
- Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues, 57*, 657-674.
- Heilman, M. E., Block, C. J., & Martell, R. F. (1995). Sex stereotypes: Do they influence perceptions of managers? *Journal of Social Behavior & Personality*.
- Heilman, M. E., Block, C. J., Martell, R. F., & Simon, M. C. (1989). Has anything changed? Current characterizations of men, women, and managers. *Journal of Applied Psychology, 74*, 935-942.
- Hersey, P., & Blanchard, K. H. (1976). Leader effectiveness and adaptability description (LEAD). *The 1976 Annual Handbook for Group Facilitators, 5*, 133-142.
- Hiller, N. J., DeChurch, L. A., Murase, T., & Doty, D. (2011). Searching for outcomes of leadership: A 25-year review. *Journal of Management, 37*, 1137-1177.
- Hoffman, C., & Hurst, N. (1990). Gender stereotypes: Perception or rationalization? *Journal of Personality and Social Psychology, 58*(2), 197-208.

- Hoffman, L. R., & Maier, N. R. (1961). Quality and acceptance of problem solutions by members of homogeneous and heterogeneous groups. *The Journal of Abnormal and Social Psychology, 62*, 401-407.
- Hoobler, J. M., Lemmon, G., & Wayne, S. J. (2011). Women's underrepresentation in upper management: New insights on a persistent problem. *Organizational Dynamics, 40*(3), 151-156.
- Hoyt, C. L. (2005). The role of leadership efficacy and stereotype activation in women's identification with leadership. *Journal of Leadership & Organizational Studies, 11*(4), 2-14.
- Hoyt, C. L., & Blascovich, J. (2007). Leadership efficacy and women leaders' responses to stereotype activation. *Group Processes & Intergroup Relations, 10*, 595-616.
- Hoyt, C. L., & Blascovich, J. (2010). The role of leadership self-efficacy and stereotype activation on cardiovascular, behavioral and self-report responses in the leadership domain. *The Leadership Quarterly, 21*, 89-103.
- Hoyt, C. L., Johnson, S. K., Murphy, S. E., & Skinnell, K. H. (2010). The impact of blatant stereotype activation and group sex-composition on female leaders. *The Leadership Quarterly, 21*, 716-732.
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60*, 581-592.
- Inzlicht, M., McKay, L., & Aronson, J. (2006). Stigma as ego depletion how being the target of prejudice affects self-control. *Psychological Science, 17*, 262-269.
- Jackman, M. R. (1994). *Velvet glove: Paternalism and conflict in gender, class, and race relations*. Los Angeles: University of California Press.

- Johns, M., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resource depletion: examining the influence of emotion regulation. *Journal of Experimental Psychology: General*, *137*(4), 691-705.
- Kalokerinos, E. K., von Hippel, C., & Zacher, H. (in press). Is stereotype threat a useful construct for organizational research and practice? *Industrial and Organizational Psychology Perspectives on Science and Practice*.
- Keller, J. (2002). Blatant stereotype threat and women's math performance: Self-handicapping as a strategic means to cope with obtrusive negative performance expectations. *Sex Roles*, *47*, 193-198.
- Keller, J., & Dauenheimer, D. (2003). Stereotype threat in the classroom: Dejection mediates the disrupting threat effect on women's math performance. *Personality and Social Psychology Bulletin*, *29*, 371-381.
- Koenig, A. M., & Eagly, A. H. (2005). Stereotype threat in men on a test of social sensitivity. *Sex Roles*, *52*, 489-496.
- Liff, S., & Ward, K. (2001). Distorted views through the glass ceiling: The construction of women's understandings of promotion and senior management positions. *Gender, Work & Organization*, *8*, 19-36.
- Lipman-Blumen, J. (1976). Toward a homosocial theory of sex roles: An explanation of the sex segregation of social institutions. *Signs*, *1*(3), 15-31.
- Logel, C., Iserman, E. C., Davies, P. G., Quinn, D. M., & Spencer, S. J. (2009). The perils of double consciousness: The role of thought suppression in stereotype threat. *Journal of Experimental Social Psychology*, *45*, 299-312.

- Lopez-Zafra, E., Garcia-Retamero, R., & Pilar Berrios Martos, M. (2012). The relationship between transformational leadership and emotional intelligence from a gendered approach. *The Psychological Record, 62*, 97-114.
- Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology, 42*, 236-243.
- Marx, D. M., & Stapel, D. A. (2006). Distinguishing stereotype threat from priming effects: On the role of the social self and threat-based concerns. *Journal of Personality and Social Psychology, 91*, 243-254.
- Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods, 44*, 1-23.
- McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology, 86*, 730-740.
- McFarland, L. A., Lev-Arey, D. M., & Ziegert, J. C. (2003). An examination of stereotype threat in a motivational context. *Human Performance, 16*, 181-205.
- McIntyre, R. B., Paulson, R. M., Taylor, C. A., Morin, A. L., & Lord, C. G. (2011). Effects of role model deservingness on overcoming performance deficits induced by stereotype threat. *European Journal of Social Psychology, 41*, 301-311.
- Miller Burke, J., & Attridge, M. (2011). Pathways to career and leadership success: Part 2—striking gender similarities among \$100 k professionals. *Journal of Workplace Behavioral Health, 26*, 207-239.

- Murphy, M. C., Steele, C. M., & Gross, J. J. (2007). Signaling threat how situational cues affect women in math, science, and engineering settings. *Psychological Science, 18*, 879-885.
- Nguyen, H.-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.
- O'Brien, L. T., & Crandall, C. S. (2003). Stereotype threat and arousal: Effects on women's math performance. *Personality and Social Psychology Bulletin, 29*, 782-789.
- O'Brien, L. T., & Hummert, M. L. (2006). Memory performance of late middle-aged adults: Contrasting self-stereotyping and stereotype threat accounts of assimilation to age stereotypes. *Social Cognition, 24*, 338-358.
- Oakley, J. G. (2000). Gender-based barriers to senior management positions: Understanding the scarcity of female CEOs. *Journal of Business Ethics, 27*, 321-334.
- Osborne, J. W. (2006). Gender, stereotype threat, and anxiety: Psychophysiological and cognitive evidence. *Journal of Research in Educational Psychology, 8*(4), 109-138.
- Oswald, D. L., & Harvey, R. D. (2000). Hostile environments, stereotype threat, and math performance among undergraduate women. *Current Psychology, 19*, 338-356.
- Paolacci, G., Chandler, J., & Ipeirotis, P. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5*, 411-419.

- Pinel, E. C. (1999). Stigma consciousness: The psychological legacy of social stereotypes. *Journal of personality and Social Psychology, 76*, 114-128.
- 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.
- Powell, G. N., Butterfield, D. A., & Parent, J. D. (2002). Gender and managerial stereotypes: Have the times changed? *Journal of Management, 28*, 177-193.
- Quinn, D. M., Kahng, S. K., & Crocker, J. (2004). Discreditable: Stigma effects of revealing a mental illness history on test performance. *Personality and Social Psychology Bulletin, 30*, 803-815.
- Rosener, J. B. (1995). *America's competitive secret: Utilizing women as a management strategy*. New York: Oxford University Press.
- Sackett, P. R., Borneman, M. J., & Connelly, B. S. (2008). High stakes testing in higher education and employment: Appraising the evidence for validity and fairness. *American Psychologist, 63*, 215-227.
- Sackett, P. R., & Lievens, F. (2008). Personnel selection. *Annual Review of Psychology, 59*, 419-450.
- Schein, V. E. (1975). Relationships between sex role stereotypes and requisite management characteristics among female managers. *Journal of Applied Psychology, 60*, 340-344.

- Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology, 94*, 168-182.
- Sekaquaptewa, D., Waldman, A., & Thompson, M. (2007). Solo status and self-construal: Being distinctive influences racial self-construal and performance apprehension in African American women. *Cultural Diversity and Ethnic Minority Psychology, 13*, 321-327.
- Shapiro, J. R., & Williams, A. M. (2012). The role of stereotype threats in undermining girls' and women's performance and interest in STEM fields. *Sex Roles, 66*, 175-183.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*, 4-28.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*, 797-811.
- Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on Black and White athletic performance. *Journal of Personality and Social Psychology, 77*, 1213-1227.
- Stone, J., & McWhinnie, C. (2008). Evidence that blatant versus subtle stereotype threat cues impact performance through dual processes. *Journal of Experimental Social Psychology, 44*, 445-452.

- Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology*, *36*, 778-793.
- U.S. Department of Labor, Bureau of Labor Statistics. (2013) *Women in the labor force: A databook*.
- Van Knippenberg, D., De Dreu, C. K. W., & Homan, A. C. (2004). Work group diversity and group performance: an integrative model and research agenda. *Journal of Applied Psychology*, *89*, 1008-1022.
- Williams, J. E., & Best, D. L. (1982). *Measuring sex stereotypes: A thirty-nation study*: Beverly Hills: Sage Publications.
- Wirth, L. (2009). *Breaking through the glass ceiling: Women in management*: Luxembourg: International Labour Organization (ILO).
- Yopyk, D. J., & Prentice, D. A. (2005). Am I an athlete or a student? Identity salience and stereotype threat in student-athletes. *Basic and Applied Social Psychology*, *27*, 329-336.
- .

APPENDIX A

NOTIFICATION STATEMENT

PROJECT TITLE: Project Test Development

INTRODUCTION

The purposes of this form are to give you information that may affect your decision whether to say YES or NO to participation in this research, and to record the consent of those who say YES. This study, Project Test Development, will take place in Mills Godwin Building, Room 222 [online].

RESEARCHERS

Principle Investigator: Debra A. Major, Ph.D., Professor, Department of Psychology
Investigator: Valerie N. Streets, B.S., Doctoral Student, Department of Psychology

DESCRIPTION OF RESEARCH STUDY

If you decide to participate, then you will join a study involving research of test development. If you decide to participate, then you will be administered a few instruments to complete on this computer. Participation will last no longer than one hour.

EXCLUSIONARY CRITERIA

You must be at least 18 years of age to complete this study.

RISKS

There are no known risks associated with participation in this study. However, as with any research, there is a possibility that you may be subject to risks that have not yet been identified.

COSTS AND PAYMENTS

The researchers want your decision about participating in this study to be absolutely voluntary. Yet if you decide to participate you will be compensated accordingly with one SONA credit [\$1.00].

NEW INFORMATION

If the researchers find new information during this study that would reasonably change your decision about participating, then they will give it to you.

CONFIDENTIALITY

Your responses in this study will be kept strictly confidential. The researchers will not attach any identifying information to your responses. The results of this study may be used in reports, presentations, and publications; but the researcher will not identify you.

WITHDRAWAL PRIVILEGE

It is OK for you to say NO. Even if you say YES now, you are free to say NO later, and walk away or withdraw from the study -- at any time. Your decision will not affect your relationship with Old Dominion University, or otherwise cause a loss of benefits to which you might otherwise be entitled. The researchers reserve the right to withdraw your participation in this study, at any time, if they observe potential problems with your continued participation.

COMPENSATION FOR ILLNESS AND INJURY

If you say YES, then your consent in this document does not waive any of your legal rights. However, in the unlikely event of harm arising from this study, neither Old Dominion University nor the researchers are able to give you any money, insurance coverage, free medical care, or any other compensation for such injury.

VOLUNTARY CONSENT

By NEXT below, you are agreeing to several things. You are saying that you have read this form or have had it read to you, that you are satisfied that you understand this form, the research study, and its risks and benefits. The researchers should have answered any questions you may have had about the research. If you have any questions later on, then the researchers should be able to answer them:

Dr. Debra A. Major
Dept of Psychology
dmajor@odu.edu
757-683-4235

Valerie N. Streets
Dept of Psychology
vstreets@odu.edu

And importantly, by clicking NEXT below, you are telling the researcher YES, that you agree to participate in this study.

Note: information in brackets indicates modifications made to the notification statement for use on Mturk.

APPENDIX B
MANIPULATION

Control Condition:

Thank you for your participation today. Old Dominion University is currently developing a series of tests to be used for students in various situations. On the following screens you will be administered one of these tests at random. You will have 10 minutes to complete that test. Please read all instructions and questions carefully and answer each question to the best of your ability.

Experimental Condition:

Thank you for your participation today. On the following screens, you will be asked to complete a test that evaluates your leadership ability. You will have 10 minutes to complete this test. Please read all instructions and questions carefully and answer each question to the best of your ability.

APPENDIX C

LEADER EFFECTIVENESS AND ADAPTABILITY DESCRIPTION (LEAD)

Directions: Assume you are involved in each of the following 12 situations. For each situation, interpret key concepts in terms of the environment or situation in which you most often think of yourself as assuming a leadership role. Say, for example, an item mentions subordinates. If you think that you engage in leadership behavior most often as a manager at work, then think about your staff as subordinates. If you feel you get the most leadership experience as an officer within a student organization, think about the members of that organization as your subordinates. If, however, you think of yourself as assuming a leadership role primarily as a parent, think about your children as your subordinates. READ each item carefully and THINK about what you would do in each circumstance. Then CIRCLE the letter of the alternative that you think would most closely describe your behavior in the situation presented. Circle only one choice. Do not change the situational frame of reference from one item to another. Select one situation as the reference for all 12 questions.

1. Your subordinates have not been responding to your friendly conversation and obvious concern for their welfare. Their performance is in a tailspin.
 - A. Emphasize the use of uniform procedures and the necessity for task accomplishment. (3)
 - B. Make yourself available for discussion but do not push. (1)
 - C. Talk with subordinates and then set goals. (2)
 - D. Be careful not to intervene. (0)

2. The observable performance of your group is increasing. You have been making sure that all members are aware of their roles and standards.
 - A. Engage in friendly interaction, but continue to make sure that all members are aware of their roles and standards. (3)
 - B. Take no definite action. (0)
 - C. Do what you can to make the group feel important and involved. (2)
 - D. Emphasize the importance of deadlines and tasks. (1)

3. Members of your group are unable to solve a problem themselves. You have normally left them alone. Group performance and interpersonal relations have been good.
 - A. Involve the group and together engage in problem solving. (2)
 - B. Let the group work it out. (1)
 - C. Act quickly and firmly to correct and redirect. (0)
 - D. Encourage the group to work on the problem and be available for discussion. (3)

4. You are considering a major change. Your subordinates have a fine record of accomplishment. They respect the need for change.
 - A. Allow group involvement in developing the change, but do not push. (2)
 - B. Announce changes and then implement them with close supervision. (0)
 - C. Allow the group to formulate its own direction. (3)
 - D. Incorporate group recommendations, but direct the change. (1)

5. The performance of your group has been dropping during the last few months. Members have been unconcerned with meeting objectives. They have continually needed reminding to do their tasks on time. Redefining roles has helped in the past.
 - A. Allow the group to formulate its own direction. (0)
 - B. Incorporate group recommendations, but see that objectives are met. (2)
 - C. Redefine goals and supervise carefully. (3)
 - D. Allow group involvement in setting goals, but do not push. (1)

6. You stepped into an efficiently run situation. The previous administrator ran a tight ship. You want to maintain a productive situation, but would like to begin humanizing the environment.
- A. Do what you can to make the group feel important and involved. (1)
 - B. Emphasize the importance of deadlines and tasks. (2)
 - C. Be careful not to intervene. (0)
 - D. Get the group involved in decision making, but see that objectives are met. (3)
7. You are considering major changes in your organizational structure. Members of the group have made suggestions about needed change. The group has demonstrated flexibility in its day-to-day operations.
- A. Define the change and supervise carefully. (0)
 - B. Acquire the group's approval on the change and allow members to organize the implementation. (3)
 - C. Be willing to make changes as recommended, but maintain control of implementation. (1)
 - D. Avoid confrontation; leave things alone. (2)
8. Group performance and interpersonal relations are good. You feel somewhat unsure about your lack of direction of the group.
- A. Leave the group alone. (3)
 - B. Discuss the situation with the group and then initiate necessary changes. (1)
 - C. Take steps to direct your subordinates toward working in a well-defined manner. (0)
 - D. Be careful of hurting boss-subordinate relations by being too directive. (2)
9. Your superior has appointed you to head a taskforce that is far overdue in making requested recommendations for change. The group is not clear about its goals. Attendance at sessions has been poor; the meetings have turned into social gatherings. Potentially, the group has the skills to help.
- A. Let the group work it out. (0)
 - B. Incorporate group recommendations, but see that objectives are met. (2)
 - C. Redefine goals and supervise carefully. (3)
 - D. Allow group involvement in setting goals, but do not push. (1)
10. Your subordinates, usually able to take responsibility, are not responding to your recent redefined of standards.
- A. Allow group involvement in redefining standards, but do not push. (2)
 - B. Redefine standards and supervise carefully. (0)
 - C. Avoid confrontation by not applying pressure. (1)
 - D. Incorporate group recommendations, but see that new standards are met. (3)
11. You have been promoted to a new position. The previous supervisor was uninvolved in the affairs of the group. The group has adequately handled its tasks and direction. Group inter-relations are good.
- A. Take steps to direct subordinates toward working in a well-defined manner. (0)
 - B. Involve subordinates in decision making and reinforce good contributions. (3)
 - C. Discuss past performance with the group and then examine the need for new practices. (1)
 - D. Continue to leave the group alone. (2)
12. Recent information indicates some internal difficulties among subordinates. The group has a remarkable record of accomplishment. Members have effectively maintained long-range goals and have worked in harmony for the past year. All are well qualified for the task.
- A. Try out your solution with subordinates and examine the need for new practices. (1)
 - B. Allow group members to work it out themselves. (3)
 - C. Act quickly and firmly to correct and redirect. (0)

D. Make yourself available for discussion, but be careful of hurting boss-subordinate relations. (2)

Note: The score for each response is indicated in parentheses. Points are awarded for each alternative action selected in response to the 12 situations. The number of points awarded is determined by how well the action selected matches the situation. Thus, a "3" response indicates the best fit. A "0" response indicates that an action was selected that has a very low probability of success. The use of a point system allows leadership adaptability and effectiveness to be expressed as a score. Possible scores range from 0 to 36.

APPENDIX D

PERCEIVED STEREOTYPE THREAT

Please think about your performance on the leadership task and respond to the following questions on a scale of 1 (not at all) to 7 (very much).

1. How often while taking this test did you think about the stereotype that women are worse at leadership tasks compared to men?
2. How much do you think that the stereotype that women are less competent leaders compared to men affected your performance on the leadership test?
3. How often did you think about performing poorly while you took the leadership test?
4. How often did you think about how members of your gender might have performed while you took the leadership test?
5. I was concerned that the researcher will judge women as a whole based on my performance on this test.
6. I worried that the researcher will think that women as a whole have less leadership ability because of how I did on this test.

APPENDIX E

EXPERIMENTER INSTRUCTIONS AND SCRIPT

Arrive to MGB 222 at least 15 minutes prior to the assigned session time. Be sure to keep the door closed until the study's start time. Check to make sure you have the most updated list of participants for the session. Log on to the appropriate number of computers. Where possible, leave a vacant station between participants to minimize distractions. For example, if there are 3 participants, log on to computers A, C, and E. If fewer than 20 participants are signed up for the session, log on to an extra computer by the door and open a web browser so participants can look up their SONA number if necessary. Ensure that you have enough SONA confirmation sheets.

At the designated start time, open the door and ask all participants to form a single-file line. One-at-a-time, ask each participant for his or her SONA number and show the participant to his or her seat based on the participant list. Instruct each participant not to touch the keyboard yet. Once all participants are seated close the door and begin introducing the study.

“Hello everyone, I would like to thank you for taking the time to participate in Project Test Development today. Your participation is very important in this research. I would like to ask that you don't touch your computer yet until we are ready to begin. First I will walk you through the collection of your SONA information. If your monitor currently appears blank you may have to shake your mouse a bit to activate it. The first screen will ask you for your SONA ID number please enter that and then press 'next.'”
On this screen you will see some questions asking about the class for which you wish to apply your credit earned for this study. Please enter that information and click 'enter.'
Once you have completed those questions please wait for me to instruct you to move on.
”

Wait until it looks as if everyone is done entering their information and then proceed.

“Has everyone answered the SONA questions? Are there any questions so far? (If there are no questions, proceed) Great, let’s begin. Click the ‘x’ in the upper right corner of your screen to close out of the SONA information page. Another window should now appear on your screen with a set of instructions. Is there anyone that does not have this screen in front of them? (Scan the room to see if anyone is having any difficulties, if so address them, if not proceed). Great, please read the instructions on this screen but do not proceed until instructed to do so.”

Allow one minute to pass. “As you progress, you will not be able to return to any previous screens. Please be sure you have read and understand the introduction to your test. When you proceed please take your time and respond to each question carefully. Even the longest experiments do not take a full hour, so you will get out of here in plenty of time; there is no need to rush. To avoid distracting others in the room, please remain seated once you have completed the experiment. Once everyone has finished, I will give you a receipt for your participation. Again, please be sure you have read the instructions on the current screen. Once you have, you may click ‘next’ and proceed.”

Sit in the back of the room and quietly observe participants to monitor their progress. Please try to avoid making noise so as not to disrupt the participants. You may use this time to enter SONA credits for the participants who are present and prep the SONA receipts. Once all monitors display the exit screen, return to the front of the room.

“I would like to thank you all again for coming today, we really appreciate your participation in Project Test Development. I am now passing out a receipt of your participation in today’s study. The top portion is for you to keep for your records should any glitches occur with the online system. Please print and sign your name on the bottom portion and return those to me. To maintain the integrity of the research we ask that you do not discuss your participation or any aspect of this study with fellow students or anyone who may be participating in this study in the future. Once you turn in your paperwork you are free to go. Thank you again for your time.”

APPENDIX F

DESCRIPTIVE STATISTICS OF OUTCOME MEASURES

Table 5.
Descriptive Statistics of Outcome Measures by Condition and Data Source

Source	LEAD			Perceived Stereotype Threat		
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
SONA	161	17.130	4.502	149	2.342	1.157
Control	77	16.377	4.152	72	2.236	1.088
Men	14	21.929	4.548	14	2.327	1.300
Women	63	15.143	2.878	58	2.214	1.042
Threat	84	17.821	4.719	77	2.440	1.217
Men	29	23.138	2.642	25	2.505	1.107
Women	55	15.018	2.725	52	2.409	1.275
Mturk	244	22.812	3.667	241	1.661	0.915
Control	127	22.969	3.528	126	1.610	0.958
Men	51	22.588	3.915	51	1.809	1.120
Women	76	23.224	3.244	75	1.475	0.811
Threat	117	22.641	3.818	115	1.716	0.865
Men	52	22.347	3.935	52	1.750	0.853
Women	65	22.877	3.735	63	1.688	0.881

VITA

VALERIE N. STREETS
5115 Hampton Boulevard
Norfolk, VA 23529

EDUCATION

- | | |
|-------------------------|--|
| Old Dominion University | Ph. D. in Industrial/Organizational Psychology
Expected, 2014
M.S. in Industrial/Organizational Psychology
Expected, 2016 |
| Towson University | B. S. with honors in Psychology and Gender Studies
May, 2011 |

AWARDS

- | | |
|-----------|--|
| 2011-2013 | Patricia W. & J. Douglas Perry Fellowship, Old Dominion University |
|-----------|--|

PUBLICATIONS

- Major, D. A., Fletcher, T. D., Streets, V. N., & Sanchez-Hucles, J. (in press). One is the loneliest number: Comparing ethnic minority solos and nonsolos. *Journal of Women and Minorities in Science and Engineering*.

BOOK CHAPTERS

- Streets, V. N., Major, D. A., & Morganson, V. J. (in press). Building inclusive IS&T work climates for women and men. In M. Khosrow-Pour (Ed.), *Encyclopedia of Information Science and Technology*, 3rd ed. Hershey, PA: Information Science Reference.
- Streets, V. N. & Major, D. A. (2014). Gender and careers: Obstacles and opportunities. In S. Kumra, R. Simpson, & R. J. Burke (Eds.), *The Oxford Handbook of Gender in Organizations*. New York: Oxford University Press.
- Streets, V. N. & Nguyen, H. (2014). Stereotype threat in the workplace. In R. Burke & D. Major (Eds.), *Men in Organizations: Allies or Adversaries to Women's Career Advancement* (pp. 270-290). Northampton, MA: Edward Elgar.
- Winstead, B. A. & Streets, V. N. (2013). Gender and workplace relationships. In H. Cooper-Thomas & R. Morrison (Eds.), *Relationships in Organizations: A Work Psychology Perspective*, (pp. 137-170). New York: Pelgrave Macmillan.